全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

译丛
"十二五"国家重点图书出版规划项目
当代财经管理名著译库

QUANTITATIVE TRADING

How to Build Your Own Algorithmic Trading Business

ERNEST P. CHAN

量化交易

如何建立自己的算法交易事业

（加）欧内斯特·陈 著
诺奇 谢彦 译
嵩 校


东北财经大学出版社
Dongbei University of Finance & Economics Press
WILEY



全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

前　言

据估计，量化交易或算法交易目前已占美国总交易规模的三分之一以上。当然，在这个领域中已有许多供机构交易员参考的关于高等数学和策略的书籍。然而，一名独立或零售交易员能够从这些算法交易中获得哪些裨益？缺少资源和运算实力的个人能够在几千种股票中进行策略回测、执行，并与实力强大的机构参与者抗衡吗？

我来告诉你如何做到这些。

目标读者

本书主要为两类读者而写：

1. 打算从事量化交易事业的独立（“零售”）交易员。

2. 有志成为大型机构量化交易员和投资经理的金融系学生或其他技术类院系的学生（本科或 MBA 水平）。

这两类截然不同的读者群体都能从相同的知识和技术中获益吗？管理 1 亿美元投资组合与管理 10 万美元投资组合有什么共同之处吗？我的观

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

点是，要想成为成功的1亿美元交易员，必先成为成功的10万美元交易员。很多事情都证明了这一点。

许多传奇的量化对冲基金经理，如普林斯顿－新港合伙公司（Poundstone，2005）的创始人Edward Thorp博士和文艺复兴科技公司的Jim Simons博士（Lux，2000），都是用自有资金开始交易生涯的。在他们开始自己的基金管理事业之前，都没在投资银行和对冲基金做过投资经理。当然，与之相反的例子也有很多，但这确实是一条比较理性的致富之路，并且对于某些富有创业激情的人来说，更是最佳选择。

即使你的目标是成为一名机构交易员，从自己的交易事业开始起步也是十分有价值的。目前的华尔街，很多物理学家和数学家都蜂拥而至。在这里，仅仅有一个名校的博士头衔是不够的。那么，通向顶级投资银行和基金的路径何在？向他们展示你拥有可以盈利的系统性方法，换句话说，就是盈利记录。拥有盈利记录，除了可以当作进入大型机构的跳板，也是独立交易员十分有价值的经历。这份经历可以使你专注于简单但盈利的策略，而不被那些深奥复杂的交易理论所牵绊，同时也可以使你专注于量化交易的本质内容，这在书本中是学不到的：比如如何在编程成本不超过1万美元的前提下建立一个指令输入系统。最重要的是，它可以使你专注于风险管理——毕竟，个人破产是有可能发生的。最后，作为在不同阶段既干过机构交易员又做过零售量化交易员的人，我只想说，要是我能在进入银行工作之前就读过这样一本书，也许我早就能够盈利了。

由于我在前言中已有所交代，所以在后面的章节中，我不会再对本书关注于自己创业的独立交易员以及他们如何建立自己的量化交易业务而进一步致歉了。同时，也希望本书能对他们走上机构交易员之路有所帮助。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

读者所需背景

虽然标题看起来有点吓人，但你完全可以直接将本书当作量化交易的指导书，而无需先成为数学或计算机领域的专家。当然，你还是需要掌握一些基本的统计学知识的，比如如何计算均值、标准差，或如何对数据点集拟合一条直线。对了，你还需要熟悉 Excel 的一些基本操作。但不需要你必须具备随机微积分、神经网络或其他高级技术。

尽管只用 Excel 就能赚到几百万美元，但还有另外一种工具可供使用。只要操作熟练，就能更有效地对策略进行回测，并且能够比使用其他工具更容易地检索和处理数据。这个工具就是 MATLAB®，它是许多机构量化交易员和投资经理都使用的数学平台。因此，我将介绍如何使用 MATLAB 来对大多数策略进行回测。我在附录中对如何使用 MATLAB 进行一些基础的编程写了一个简要的辅导材料。对于许多零售交易员来说，MATLAB 的价格太过昂贵，不过第 3 章“回测”中会介绍一些相对便宜的替代软件。并且，许多大学生可以购买比较便宜的 MATLAB 学生版本，也可以在学校里免费使用。

本书内容

本书绝不是一本量化交易技术或量化交易术语的百科全书，也不是专门介绍一些特殊的盈利策略（尽管你可以将书中所举例的少数策略进一步完善以获得更高的收益率）。相反，这是一本教你如何自己去寻找盈利策略的书。它会告诉你优等策略的特征是什么，如何通过对一条策略进行优化和回测来确认其是否具有良好的历史业绩，并且最重要的是，确认这条

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

策略在未来能否使你继续盈利。本书也教你根据策略的真实盈利性来调整交易规模的系统方法。还教你如何在家中构建自动交易执行系统的具体细节。最后，本书教你一些风险管理的基础知识。想在长期中生存下来，知道如何进行风险管理是非常重要的。如果你想享受交易员人生（不只是盈利），还需要躲避一些心理陷阱。

尽管寻找优等策略的基本技术对任何流通证券都适用，但我所举的例子都集中在我最熟悉的交易领域：股票的统计套利交易。“回测”一章讨论了股票、期货、外汇的历史数据来源，但并不包含期权，因为期权并不在我擅长的研究范围内。

本书内容结构基本上是按照交易员开展量化交易事业的各项步骤顺序来组织的。首先从寻找切实可行的策略开始（第 2 章），接下来对策略进行回测，确保其至少在历史上有良好的业绩（第 3 章），并构建业务和技术的基础设施（第 4 章），建立执行策略的自动交易系统（第 5 章），以及对策略所生成的持有头寸进行资金和风险管理（第 6 章）。我将在第 7 章讲述一些资深交易员所熟知的重要前沿概念，最后在第 8 章对独立交易员如何寻找适合自己的位置以及如何扩大交易事业做一些思考。同时我写了一章附录，对如何使用 MATLAB®提供指导。

本书还包括两类方框材料：

- 阐释或图解一项概念的工具条
- MATLAB 或 Excel 的代码举例

如果读者希望学习更多内容，并且想掌握量化交易的最新新闻、理念和发展趋势，欢迎访问我的博客（epchan. blogspot. com）和提供高级内容的网站（pchan. com/ subscriptions），我会尽我所能回答大家的问题。我的高级内容网站上有不少前沿文章以及若干盈利策略的回测结果。本书读者可以免费访问高级内容网站，在后面的章节中可以找到进入网页的密码。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

致　谢

在投资银行（摩根士丹利、瑞信、Maple Securities）和对冲基金（Mapleridge Capital、Millennium Partners、MANE Fund Management）工作时，我的同事和导师向我传授了许多量化交易的知识和经验，谨在此对他们多年来给予的建议、指导和帮助表示衷心的感谢。成为独立交易员和独立顾问以来，我从客户、博客读者、博友和许多合作对象那里受益良多。在这里，向 Steve Halpern 和 Ramon Cummins 致谢，他们审阅了本书初稿并指正了部分错误内容；感谢 John Rigg，对我博客中的部分文章提出了建议，很多博客内容在本书中都有体现；同时也要感谢 Ashton Dorkins（tradingmarkets. com 网站的首席编辑）帮助发表了我的博客文章；还要向 Yaser Anwar 致谢，将本书向其知名投资博客的读者做了大量宣传。感谢 John Wiley & Sons 出版社的编辑 Bill Falloon 对本书提出的建议，策划编辑 Emilie Herman、出版编辑 Christina Verigan 为本书出版发行所做的辛勤工作。最后，感谢 Ben Xie 所要求的写作原则：简单的就是最好的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

目　录

第 1 章　量化交易初探　1

1.1　谁能成为量化交易员？　2

1.2　量化交易的特点　4

1.3　漫漫前路　7

第 2 章　寻找切实可行的策略　8

2.1　甄别适合自己的策略　11

2.2　识别貌似可行的策略及其陷阱　15

2.3　小　结　24

第 3 章　回　测　26

3.1　常用的回测平台　26

3.2　查找与使用历史数据库　31

3.3　业绩度量　37

3.4　避免常见的回测陷阱　44

3.5　交易成本　54

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

3.6 策略改进 59
3.7 小 结 60

第4章 创建交易业务 62
4.1 业务结构：零售还是自营？ 62
4.2 选择一家零售经纪公司（或自营交易公司） 64
4.3 设 备 67
4.4 小 结 68

第5章 交易执行系统 70
5.1 自动交易系统的功能 70
5.2 最小化交易成本 77
5.3 用仿真交易测试交易系统 78
5.4 实际业绩偏离预期的原因 79
5.5 小 结 81

第6章 资金和风险管理 83
6.1 最优资本配置和杠杆 83
6.2 风险管理 92
6.3 做好心理准备 96
6.4 小 结 99
6.5 附录：收益率正态分布时凯利公式的简单推导 99

第7章 量化交易专题 101
7.1 均值回归策略和惯性策略 101
7.2 状态转换 105
7.3 平稳性和协整性 111

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.4 因子模型 118
7.5 清仓策略 124
7.6 季节性交易策略 128
7.7 高频交易策略 137
7.8 高杠杆组合优于高贝塔组合吗？ 139
7.9 小 结 140

第 8 章 结语：独立交易员能否成功？ 142
8.1 接下来 145

附录：MATLAB 快速回顾 147

参考文献 154

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第1章 量化交易初探

如果你怀着好奇心拿起此书，可能已经听说过量化交易。对于那些在各大主流媒体上听说过这种交易的读者而言，仍不乏一些常见的错误观念，值得厘清。

量化交易，也称算法交易，是严格按照计算机算法程序给出的买卖决策进行的证券交易。通过历史金融数据检验编码策略历史业绩的基础上，交易员设计甚至亲自编写计算机算法程序。

那么，量化交易仅仅是技术分析的一个花哨名字么？可以说，基于技术分析的策略如果能够完全编码为计算机程序，就可以作为量化交易系统的一部分。然而，并非所有技术分析都能被认定为量化交易。例如，某些图表分析技术，如“寻找头肩形态”，可能就不在量化交易员的考虑范围，因为它们相当主观，并且很难量化。

所以，量化交易所包含的内容要比技术分析丰富得多。许多量化交易系统在进行运算输入时会用到基本面数据，例如收入、现金流、权益负债率以及其他数据。毕竟，基本面数据仅仅是数字而已，而计算机完全能够处理任何输入的数字。当需要对一家公司目前的财务业绩进行横向或纵向比较时，计算机通常能够做到与人工分析一样出色，而且计算机还可以同

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

时跟踪数千家公司。一些先进的量化交易系统甚至可以将新闻事件转换成输入变量。如今，使用计算机来解析和解读新闻报道已经成为可能。（笔者就职于 IBM 时就曾涉足这一领域，研究计算机系统如何粗略读取文档内容的问题。）

所以我们或许可以这么说：只要能将信息转换成计算机能够读懂的比特和字节，就能将其看作是量化交易的一部分。

1.1 谁能成为量化交易员？

一个事实是，绝大多数机构量化交易员拥有物理、数学、工程或计算机科学的高学历。分析或交易复杂的衍生工具，这些自然科学的训练通常是必不可少的。但是衍生工具并不是本书关注的重点，也不是说只有通过复杂的金融工具才能够赚钱（事实上，交易复杂的住房抵押贷款支持证券就可能变得一贫如洗，正如 2007—2008 年金融危机以及贝尔斯登的倒闭）。本书所讨论的量化交易类别称作“统计套利交易”。统计套利交易所处理的都是最简单的金融工具：股票、期货或外汇。想要成为一名统计套利交易员，不需要很高的学历。只要上过高中的数学、统计学、计算机编程和经济学课程，就具备处理某些基本统计套利策略的能力了。

看到这里，你知道了进行统计套利交易并非一定需要高学历，但你也许会问，假若拥有高学历肯定会带来某些优势吧？未必。笔者曾获得世界顶尖物理学院的博士学位（不妨直说，康奈尔大学），曾是世界级的顶尖计算器科学研究团队（高科技的殿堂：IBM 的托马斯·约翰·沃森研究中心）的明星研究员。之后在许多顶级投资银行和对冲基金担任研究员以及交易员，包括摩根士丹利、瑞信等。在这些权威机构担任研究员和交易员时，我总是努力尝试使用一些学过的高等数学知识和技术，并将其运用到统计套利交易中。可在进行了上亿美元的交易后收获了什么呢？只是给

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

我的雇主和投资者造成了亏损、更多的亏损、一眼望不到头的亏损。最后，我灰头土脸地离开了金融行业，在家搭建了一个简易卧室作为我的交易办公室，开始进行一些最简单的量化策略交易。这些策略是任何一个头脑聪明的高中生都能轻而易举搞定的。那是我人生中，第一次建立起能获利的策略（例 3.6 描述了其中一个策略），此后亦如此。我的经验教训是什么呢？正如爱因斯坦所说的："任何事情都应该使其尽可能地简单，直到无法再简单为止"。

（第 8 章还会更详细说明，为什么独立交易者能够打败大型投资机构的基金经理。）

我成为量化交易员走的是一条比较传统的途径，但很多人不是。那么，典型的独立量化交易员都是些什么人呢？就我所知，他们包括已倒闭对冲基金的前交易员，经纪公司前计算机程序员，交易所的前交易员，前投资银行家，前生物化学家，以及前建筑师等。有些人接受过高级技术培训，但其他人只有高中水平的统计学知识。他们中的多数人使用 Excel 之类的基本工具来进行回测，尽管有些人也许会聘请程序员来帮忙。多数人曾在其职业生涯的某段时期与金融界有过接触，但现在他们都觉得成为一名独立交易员更适合。据我所知，他们的业绩大都非常出色，同时还能尽情享受独立工作所带来的充分自由。

除了在之前的工作岗位上涉猎了不少金融知识外，一个不容忽视的事实是，这些交易员在那时的收入让他们有本钱独立创业。当一个人决定投身于独立交易后，对孤立无助的害怕和对交易亏损的畏惧是在所难免的。因此，能够对风险进行提前预判及有足够的存款是很有帮助的。同时，不需要用交易的收益来维持日常生活也是非常重要的前提，因为并不是很快就可以找到能够获得稳定收益率的策略（见第 6 章）。

与某些人对此感到恐惧不同，另一些人投身于此的原因是对惊险和刺激的偏爱，以及对能够迅速获利的超强自信。这同时也是一种会带入独立量化交易中的危险情绪。正如我希望在本书中劝告你的，迅速获利并不是

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

量化交易的目的。

因此，理想的量化交易员应是这样的人：之前有一定的金融或计算机编程的经历，有足够的存款能够用来应对不可避免的亏损和收入空窗期，并且能够在贪婪和恐惧的情绪中找到恰当的平衡。

1.2 量化交易的特点

我们中的很多人出于各种原因从事量化交易，如令人兴奋、挑战智力、增加财富，甚至只是因为这是我们唯一擅长的事情。但对于还有其他技能和机会的人来说，量化交易是否是最佳职业选择，还是要深思熟虑。

无论把对冲基金的巨大规模和数以亿计的巨额财富传说得如何神乎其神，在许多情况下，开始量化交易生意与开始任何小生意一样，都是从一点一滴起步的。我们需要从小规模、有限额（也许初始投资只有 5 万美元）的投资入手，在获得经验及有盈利之后再逐步增加交易额度。

然而从其他方面讲，量化交易生意又与其他小生意很不一样，在此列出一些重要的区别。

1.2.1 易扩大

与绝大部分生意相比（办网站除外），量化交易的规模是很容易改变的（一定程度上）。只要策略能够持续盈利，就可以在家里轻易操作百万美元等级的交易。这是因为扩大交易规模通常只是修改交易程序中一个参数而已。这个参数叫做“杠杆”。你无需与银行家或风险投资家谈判以获取更多的交易资金，经纪人通常会乐于效劳。如果你是自营交易公司的成员（第 4 章会详细讨论这个问题），甚至能够获得超过证券交易委员会（SEC）规则 T 所允许的杠杆上限。即使账户中只有 5 万美元股票，自营交易公司允许你在一天之内进行 200 万美元投资组合的交易也是常有的事

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

(40 倍杠杆)。同时，量化交易绝不是一条快速致富的捷径。你应该寄希望于收益的稳定增长，但基本不会像创办一个网站或一家软件公司那样，一年获利 200%。事实上，正如我将在第 6 章“资金和风险管理”中所阐述的，为了追求短期收益而过度杠杆化是很危险的。

1.2.2 节省时间

从事大多数的小生意都很耗费时间，至少在其初始阶段是这样，而量化交易所花费的时间则相对较少。因为量化交易本质上是一种高度自动化的生意。有时，越是人为干涉系统程序、修改决策，业绩可能反而越差(详见第 6 章)。

每天需要在量化交易上花费多少时间，完全取决于所能达到的自动化程度。例如，在我曾工作过的一家对冲基金，有些同事一个月才去一趟办公室。其余时间，他们只是偶尔在家远程监控一下正进行交易的办公室计算机。

就我自己而言，自动化程度处于中游水平。我花费时间最多的时候就是每天早晨开盘之前：我通常打开不同的程序来下载和处理最新的历史数据，阅读提醒屏上的公司新闻，运行程序生成当日的交易下单指令，然后在开盘前下达一揽子新的指令，并启动当日自动下单程序。同时我会在电子表格中更新前一交易日不同策略的盈亏记录。这些事全部做完大概需要两个小时。

之后，我还会在收盘前花半小时手动执行平仓程序，检查清仓指令执行状况，并适时关闭各种自动交易程序。

在每日的开盘时间，所有事情都是自动运行的。当然，有时也会事与愿违：我常常忍不住要去看一眼（有时是很多眼）交易屏幕，看看不同的策略在当天的盈亏情况。极端情况下，盈亏的剧烈波动使我感到恐惧，会有一种立刻手动清仓的冲动。幸运的是，随着时间推移，我比以前更能抑制住这种冲动。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

当我有大把的空闲时间时，这种人为干涉的冲动也很强烈。因此，与其盯着交易屏幕，倒不如把注意力转移到其他更健康、更有趣的活动上，比如在交易日去健身房锻炼。

我说量化交易不怎么耗费时间是指它的操作方面。如果想要发展你的业务，或者想在日益激烈的竞争中保持当前的收益，就需要花时间对新的策略进行研究和回测。对任何事业来说，研究和开发都是必要的创新工作，不过何时该进行这些工作，就比较自由、因人而异了。所以，在开盘时间，我搞研究、回复邮件，与其他交易员、研究员或客户聊天，去健身房锻炼，等等。有时我也在晚上或周末做这些事情，但仅仅是因为我想去做，而不是因为我不得不做。

当我的收益提升时，就会增加对软件开发的投入，以提高自动化程度。交易程序能在恰当的时间自动开始运行，自动下载数据，甚至能自动解读新闻并采取适当回应，以及在收盘后自动关闭，这些都是我期望改进的方向。如果这一天能够到来，日常操作也许不用花费任何时间，哪怕我在度假，交易程序也可以像往常一样自动运行，当出现问题时，交易程序能够通过我的手机或技术支持服务发出警报。简而言之，如果你想拥有更多的休闲时间，或者把时间和财务资源用在开拓其他业务上，量化交易就是适合你的工作。

1.2.3 营销非必需

量化交易与其他小生意有个最大、最明显的不同：对于绝大多数小生意来说，营销起着决定性作用，毕竟你要从其他人手里赚钱，而他们是要根据很多因素做出购买决定，并不仅仅看价格；在量化交易中，金融市场上的交易对手“仅仅”根据价格作出购买决定。除非你是在为他人管理资金（这已超出本书讨论范围），在量化交易业务中绝对不需要营销。这一点似乎显而易见且微不足道，但这的确是一个很重要的不同之处，因为量化交易只要求专注于自己的产品（策略和交易软件），而无需受他人看法

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的影响去做任何事。对许多人而言，这也许是量化交易生意的终极魅力。

1.3 漫漫前路

在你确信想要从事量化交易后，一大堆问题便会接踵而来：如何找出合适的策略？如何在进行回测前就能辨别出一个策略的优劣？如何对这些策略进行严密的回测？如果回测业绩不错，在交易结构和技术设备方面，你需要采取哪些措施来执行策略？如果这项策略在前期的真实交易中获利，该如何扩大规模以增加收入，同时又能管理好那些偶尔出现，但只要是交易就难以避免的亏损？第 2 章到第 6 章将解答这些问题。

我们的终极目标是为了取得能够长期持续增长的盈利。尽管为了达到这个目标可能需要一个艰苦而漫长的过程，但事实上，与其他许多小生意相比，这个过程已是相对短暂且容易的了。在我刚成为独立交易员时，仅仅用了三个月便找到了我的第一个策略并对其进行了回测，开立了一个 10 万美元的经纪账户，执行交易程序并根据策略开展交易。很快，策略在第一个月便开始盈利。之前在互联网创业时期，我开办过一家软件公司，却花费了比现在多 3 倍的投资、5 倍的人力和 24 倍的时间证明了该商业模式的失败，使得包括我在内的所有投资者血本无归。与那段经历相比，量化交易可爱多了，最重要的是，它真的能赚钱。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 2 章　寻找切实可行的策略

不用奇怪，寻找交易理念事实上“并不是”打造量化交易业务的最困难部分。我们每时每刻都能廉价甚至免费地从公开空间找到成百上千项策略，并且许多策略拥有者都会愿意告之完整的交易方法和回测结果。我们可以从金融投资方面的相关书籍、报纸杂志、主流媒体网站、网上或附近公共图书馆的学术论文、交易员论坛、博客等处获得这些策略。表 2－1 列出了笔者认为有价值的部分资料来源，当然，这些都只不过是众多资料中的沧海一粟。

以前，出于对学术的偏好，我会定期阅读许多商学院教授的预印本，或下载最新的金融期刊文章，以寻找那些潜在的好策略。事实上，在开始独立交易时，我的第一个策略（第 7 章中 PEAD 策略的一个版本）的灵感就源于学术研究。不过后来，我很快发现学术领域提出的很多策略不是过于复杂，已经过时（曾经盈利的策略可能由于竞争已经失效），就是在其进行回测时所需的数据过于昂贵（如历史的基本面数据）。另外，许多这类学术策略只对小盘股有效，而小盘股较差的流动性，会导致实际交易盈利远不如回测结果那么好。

当然，如果你足够有耐心，还是能有所发现的。但我发现在交易员论

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

表 2—1 **交易理念的来源**

类型	网址
学术	
商学院金融学教授的网站	www. hbs. edu/research/research. html
社会科学研究网络	www. ssrn. com
联邦经济研究局	www. nber. org
商学院的量化金融研讨会	www. ieor. columbia. edu/seminars/ financia lengineering
《纽约时报》周日商业版 Mark Hulbet 专栏	www. nytimes. com
《经济学家》杂志金融版 Buttonwood 专栏	www. economist. com
金融网站和博客	
雅虎财经	finance. yahoo. com
交易市场	www. TradingMarkets. com
寻找阿尔法	www. SeekingAlpha. com
街道网	www. TheStreet. com
科克报告	www. TheKirkReport. com
阿列亚的博客	www. aleablog. com
异常收益率	www. AbnormalReturns. com
Brett Steenbarger 交易心理学	www. brettsteenbarger. com
本书作者的博客	epcha. blogspot. com
交易员论坛	
精英交易员	www. Elitetrader. com
财富实验室	www. wealth-lab. com
报纸杂志	
股票、期货和期权杂志	www. sfomag. com

坛或博客里，能够发现一些更简单且同样可以盈利的策略。你或许不相信，怎么可能会有人真的愿意把那些能够盈利的策略放在公共空间无偿分享给他人。毕竟，这种行为会加剧竞争并且降低策略的盈利能力。这么想

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

是有道理的，因为在这些空间所找到的很多现成的策略，都经不起严格的回测。和学术研究成果一样，来自交易员论坛的策略，要么只在有限时间内有效，要么只对特定种类股票有效，要么只在忽略交易成本条件下有效。不过，真正的窍门是：对基础策略进行变形，并用于赚钱。（第 3 章将详细介绍某些基础策略及其常见变形。）

例如，有人曾向我推荐财富实验室（见表 2－1）论坛上的一项策略，据称具有高夏普比率。但当我对其进行回测后发现，它完全没有声称的那样出色。于是我开始对这项策略进行一些简单的调整，如缩短持有期、改变建仓和清仓的时点，最终这一策略成了我主要的盈利来源。如果你足够勤奋且富有创新精神，能够对基本策略进行各种变形，肯定可以从中发现高盈利的策略。

当我离开资产管理机构成为一名自由交易员时，对自己无法再与同事和导师交流交易理念而忧心忡忡。但我随后发现，获取与分享交易理念的最好途径之一，是建立自己的交易博客。把自己发现的交易“秘密”通过博客与他人分享，你会从读者那里获得更多的回赠。（那个向我推荐财富实验室策略的朋友就是我博客的一名读者，如果不是因为博客，相隔 12 个时区的我们根本毫无交集，我也就无法从他的建议中获益。）其实，那些你认为是秘密的策略多半也早已为他人所知。一项策略真正的独有价值和值得保密的地方是你自己的窍门和所进行的变形，而绝不是基础版本。

另外，网上评论能及时筛选掉那些劣等策略，从而避免较大亏损的发生。我曾经在博客上极力推荐由几个金融学教授提出的一个季节性股票交易策略，很快有个读者就通过回测否定了这一策略（可参见我博客中的《股票的季节性交易》及读者评论，epchan. blogspot. com/2007/11/seasonal-trades-in-stocks. html。有关此策略的详细介绍，见例 7.6。）当然，无论如何我都不会在没有做过回测的情况下，就对一项策略进行实盘交易。事实上，我之后对这项策略所进行的回测也证明了那个读者的结论是对的。其实，他所发现的策略的重大缺陷也正是对我的回测结果正确的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

有力证明。

总之，相比于为纽约的神秘对冲基金工作，我觉得作为一名独立交易员实际上更容易收集和交流策略。（我在一家位于第五大道、管理着数十亿美元的对冲基金 Millennium Partners 工作时，亲眼看到一名交易员从他的程序员手中抢过一篇“公开发表”的论文，其实这名程序员只是碰巧从交易员的桌子上拿起来而已。他担心那个程序员知道他的“秘密”。）也许只有当我们觉得就算有人用 1 亿美元进行模仿交易也不会让我们变得无利可图时，才不会太担心自己的秘密被他人知晓。

对于我们而言，真正困难的地方并不是缺乏交易理念，而是缺乏甄别策略的能力。这种甄别能力需要我们判断一项策略是否适合自己的实际情况和交易目标，需要在花费大量时间进行回测之前就能判断出策略是否可行。本章主要讨论如何进行策略的甄别。

2.1 甄别适合自己的策略

一项策略可行与否通常并不取决于策略本身，而取决于使用策略的人。下面是一些需要考虑的因素。

2.1.1 工作时间

交易是否只是你的兼职工作？如果是，那你或许只能考虑那些隔夜持仓的策略，而不是频繁进行日内交易的策略。否则的话，你就必须考虑使用那些完全自动交易的策略（见第 5 章），这些策略能在绝大多数时间里自动交易并在出现问题的时候发出警告。

当我还在全职工作，只能利用业余时间给自己做些交易的时候，个人账户采用一项简单的策略，只需在每次开市前，下达或者调整少数交易型开放式指数基金（ETFs）的限价指令。后来，我成为了一名独立交易员，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

由于自动化水平还相对较低，所以只考虑那些仅仅要求在开市前和闭市前各下一次指令的策略。之后，我增加了一个程序，这个程序能自动实时获取市场数据，并在某些条件被触发时，在交易日自动给经纪账户下指令。于是，交易依然只占据我生活的一小部分时间。这也是我进行量化交易的首要原因。

2.1.2 编程水平

你善于编程吗？如果你会 VB、Java、C#或C++等编程语言，就可以开发高频交易策略，交易大量的证券。否则的话，你就应该选取那些每天只交易一次或只交易少数股票、期货或外汇的策略。（聘请软件工程师可以解决这一问题。第 5 章会详细说明。）

2.1.3 交易资本

你是否有足够的资本可用于交易，以及支付设备和运营费用？我并不推荐资本规模少于 5 万美元的账户采用量化交易策略进行交易。我们以 10 万美元作为大资本账户和小资本账户的分界线。可用资本会影响许多选择，首先就是应该开立零售经纪账户还是自营交易账户（详见第 4 章）。目前，我们要清楚的是，可用资本是策略选择的一个制约条件。

对于小资本账户，就需要寻找一个可以最大限度发挥杠杆作用的策略（当然，只有在策略本身可以盈利的情况下，使用更高的杠杆才能获得更多的收益）。交易期货、外汇和期权的杠杆要比股票高；规则 T 允许日内交易头寸可以有 4 倍杠杆，而日间头寸只有 2 倍杠杆。最后，可用资本（或杠杆）决定了你应该进行单向交易（纯多头或纯空头）还是货币中性交易（对冲交易或配对交易）。一个货币中性组合（多头头寸的市场价值等于空头头寸的市场价值）或市场中性组合（投资组合的贝塔值接近于 0，这里贝塔值是投资组合的期望收益率与市场组合期望收益率之比），需要的资本或杠杆是单向组合的两倍。此外，虽然对冲头寸比非对冲头寸的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

风险要小，但收益率也相应更低，可能无法满足你的个人需求。

可得资本也会带来一些间接约束。它会制约你在设备、数据库和软件上的花费。例如，如果交易资本较少，你的在线经纪商就不大可能愿意向你提供很多股票的实时市场数据，你也就无法采用需要覆盖众多股票实时市场数据的策略。（当然，你可以向第三方数据供应商订购数据，但考虑到交易资本较少，这种额外的成本可能并不合适。）同样，准确的高频历史数据比日历史数据昂贵，所以，高频股票交易策略对于资本量低的账户并不适用。历史数据还有一个比频率更重要的特征要考虑：数据是否有生存选样偏差。下节会给出存活偏差的定义。这里，我们只需要知道，无存活偏差的数据要比包含有选样偏差的数据昂贵很多。而且，如果数据有存活偏差，回测的结果很可能是不可靠的。

同样的考虑也适用于新闻，能否支付得起像彭博资讯这样覆盖面广、提供实时新闻的数据终端，决定了你能否采用新闻驱动型策略。同样，能否支付得起拥有完整历史基本面数据（如公司财务数据）的数据库，决定了你能否采用依赖这些数据的策略。

表2—2罗列了资本（包括用于交易和各种支出）约束如何影响选择。

表2—2 **可得资本如何影响你的选择**

小资本账户	大资本账户
自营交易公司会员	零售经纪账户
期货、外汇及期权	所有证券，包括股票
日内交易	日内交易及日间（隔夜）交易
单向交易	单向交易或市场中性交易
日内交易少数股票	日内交易大量股票
有存活偏差的日历史数据	无存活偏差的高频历史数据
覆盖面窄的延迟新闻来源	覆盖面广的实时信息来源
没有历史新闻数据库	无存活偏差的历史新闻数据库
没有股票历史基本面数据	无存活偏差的股票历史基本面数据

当然，此表所列的并不是硬性规定，仅仅是一些需要考虑的因素。例

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

如，如果你的资金规模较小但却在一家自营交易公司开户，那表中的许多问题就可以不用考虑（设备支出除外）。在我刚开始独立量化交易员的生涯时，只在零售经纪公司（我选择了 Interactive Brokers）有一个 10 万美元的账户，从而采用单向的日内股票交易策略。当我开发了一个有时需要更多杠杆来实现盈利的策略时，我同时成为自营交易公司的会员。(是的，你可以同时拥有两个甚至更多的账户。事实上，仅仅是为了执行速度和流动性获得的不同，就值得这么做。见第 4 章的“选择一家零售经纪公司(或自营交易公司)”。)

尽管我在本书中一直警告要注意历史数据的存活偏差，但我刚开始使用 HQuotes. com（第 3 章还会介绍更多的数据库和工具）的下载程序，从雅虎财经下载的是分拆和股息调整后的金融数据。这一数据库并不是无存活偏差的，但在之后的两年多，我一直主要使用这一数据库进行回测！事实上，我认识的一名交易员，他每天的交易量是我账户的 10 倍以上，也主要用这一有偏差的数据进行回测，而且他的策略还是盈利的。为什么会这样？或许是因为采用的是日内交易策略。而那些愿意并且有能力购买无存活偏差数据的交易员，多是那些交易量达数千万美元或更多的资金管理机构的从业者（包括之前的我）。由此你可以明白，只要能认识到工具和数据的局限性，你还是可以通过各种办法取得成功的。

尽管期货提供了高杠杆，但很多期货合约规模很大，小资本账户仍然无法交易。比如，纽约商业交易所（NYMEX）的铂金期货合约只需要 8 100美元的保证金，但面值高达 10 万美元。另外，该合约 6%的日波动率很常见，这意味着仅仅一份合约，账户的日盈亏波动为 6 000 美元（相信我，我过去曾在个人账户中交易过这类合约，它的价格变化与所希望的方向背道而驰时总是让我倍感头疼）。相反，芝加哥商品期货交易所(CME，已与纽约商业交易所合并）的 E-迷你标准普尔 500 指数期货(ES）的面值为 67 500 美元，因其日内超过 6%的波动 15 年来只发生过 2 次，所以只需要 4 500 美元保证金，仅有铂金期货合约的 55%。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2.1.4 目标

大部分选择成为交易员的人，都希望能够获得稳定（最好是增长）的月度收入或季度收入。但是如果你财富充裕，长期资本收益才是最重要的。策略上是追求长期资本收益还是短期收入，主要根据持有期来区分。显然，如果你持有一只股票平均长达一年，就不大可能会有很高的月度收入（除非从一开始你就每个月构建一个新的子组合，且持有一年。也就是说，分时入市）。更微妙的是，尽管根据策略，持有一只股票的平均期限只有一个月，月度之间的收益率波动可能会很大（除非你的投资组合包含数百只股票，这也是分时入市的一个结果），以至于无法使用月度收入来评估业绩。持有期（或相反，交易频率）和收益持续性（即夏普比率，或相反，挫跌）之间的关系，我们将在接下来的部分进行更深入的讨论。简而言之，如果希望更频繁地实现盈利或赚取收入，你的持有期就应该尽可能地短。

有些投资顾问散布一种错误的观点：如果你的目标是最大化长期资本增长，最好的策略是买入并长期持有。这一说法在数学上早已被证明是错误的。实际上，“只要你能获得足够高的杠杆”，最大化长期资本增长可以通过最大化夏普比率（下节会给出定义）的策略实现。因此，即使你的目标是长期增长，若不考虑税收以及保证金借贷限制，在一个持有期很短、年度收益率较低、夏普比率很高的短期策略，与一个持有期很长、年度收益率较高、夏普比率较低的长期策略之间，你依然应该选择短期策略（在第 6 章“资金和风险管理”中会详细讨论这一令人吃惊的事实）。

2.2 识别貌似可行的策略及其陷阱

现在，假设你找到了一些可能符合你个人要求的策略，并且已经有人

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

对这些策略做过回测，称其具有很高的历史收益率。在你着手对这些策略进行完整的回测之前（更不用说投入资本进行实盘交易了），可以进行一些快速测试方法，以节省时间和成本。

2.2.1 策略与基准相比如何？收益持续性如何？

这一点对于一个买入（而非卖空）股票的策略而言，似乎是显而易见的。大家可能都知道，如果一个纯多头策略的年收益率为10%，并没什么吸引力。因为平均而言，投资一个指数基金的年收益率就算没有更高也至少会有这个水平。但是，对一个货币中性策略而言（持有相同价值的多头和空头），10%则是非常高的收益率了，因为这一策略的比较基准不是市场指数，而是无风险资产，如3个月期美国国债（本书写作时在4%左右）。

此外，还要考虑策略的获利持续性。也许一个策略与基准有相同的平均收益率，但在基准偶尔下跌的月份里，每月仍然获得正的收益。这时，我们认为这一策略是优于基准的。这也促使我们使用信息比率或夏普比率（Sharpe，1994），而非收益率作为量化交易策略的业绩衡量指标。

信息比率是评价一个多头策略时所用的衡量指标。定义如下：

信息比率＝超额收益率的均值/超额收益率的标准差

其中：

超额收益率＝组合收益率－基准收益率

基准通常是指所交易证券所属的市场指数。例如，如果只交易小盘股，对应的市场指数就是标准普尔600小盘股指数或罗素2000指数，而不是标准普尔500指数。如果交易的是黄金期货，对应的市场指数是黄金的现货价格，而不是股票指数。

夏普比率是信息比率的一个特例，适用于货币中性策略。此时，通常选择无风险利率作为基准。在实践中，大多数交易员在使用单向策略（纯多头或纯空头）时，也会选用夏普比率作为衡量指标，因为能够对不同种

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

类的策略进行比较。大家对无风险利率比较有共识，但在市场指数的选择上，不同的交易员有不同的看法。这就使得不同交易员之间的信息比率不具备可比性。

（实际上，在计算夏普比率时如何减去无风险利率、如何年化夏普比率以用于对比等细节问题是有一些技巧的。下一章会介绍这些技巧，还会给出计算货币中性策略和纯多头策略的夏普比率的例子。）

你肯定很好奇，如果夏普比率真的能这么出色地对不同策略进行比较，为什么我们在实际工作中还是主要使用收益率作为策略业绩的衡量指标呢？有一次，我和一名同事去 SAC 资本顾问公司（资产管理额达 140 亿美元）推广一项策略，其风险管理部门的负责人对我们讲："当然，夏普比率高确实很好，但是只有获得更高的收益率，我们才能够拿到更多的奖金。"这种观点是非常错误的：越高的夏普比率事实上使得你最终获利越多，因为高夏普比率让你可以运用更高的杠杆进行交易。最终，重要的是策略的杠杆收益率，而不是名义收益率。第 6 章"资金和风险管理"会对此作进一步阐释。

（那次面向 SAC 资本顾问公司的推广并不成功，并不是因为策略的收益率不够高，而是那时我和我的同事对夏普比率和杠杆收益率之间的数学关系还不熟悉，以至于没能很好地说服该风险管理部门的负责人。）

现在，你已经知道了什么是夏普比率，你也许还想知道备选策略的夏普比率的大小。策略的作者通常不会公布这些数字，但你可以私下发电子邮件向他们索取。很多时候，这些作者（尤其是金融学教授）都会乐意提供这些数据，但如果被拒绝，除了自己进行回测外，也别无选择了。不过，有时候你仍然可以在信息不足的情况下，对夏普比率做出有根据的猜测：

- 如果一项策略的年交易次数有限，夏普比率很可能就不会太高，但这并不影响你把这项策略作为多元策略的组成部分，但不足以把它当作主要盈利来源的策略。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

• 如果一项策略的挫跌很大（如超过 10%），或是挫跌时间很长（如 4 个月或更长），也不大可能有很高的夏普比率。下一节会详细介绍什么是挫跌，但你可以直接观察净值曲线（在没有撤资和注资的条件下，也即累计盈亏曲线）是否有很大的起伏。曲线上任意峰尖到谷底的部分都是一个挫跌（见图 2—1 的例子）。

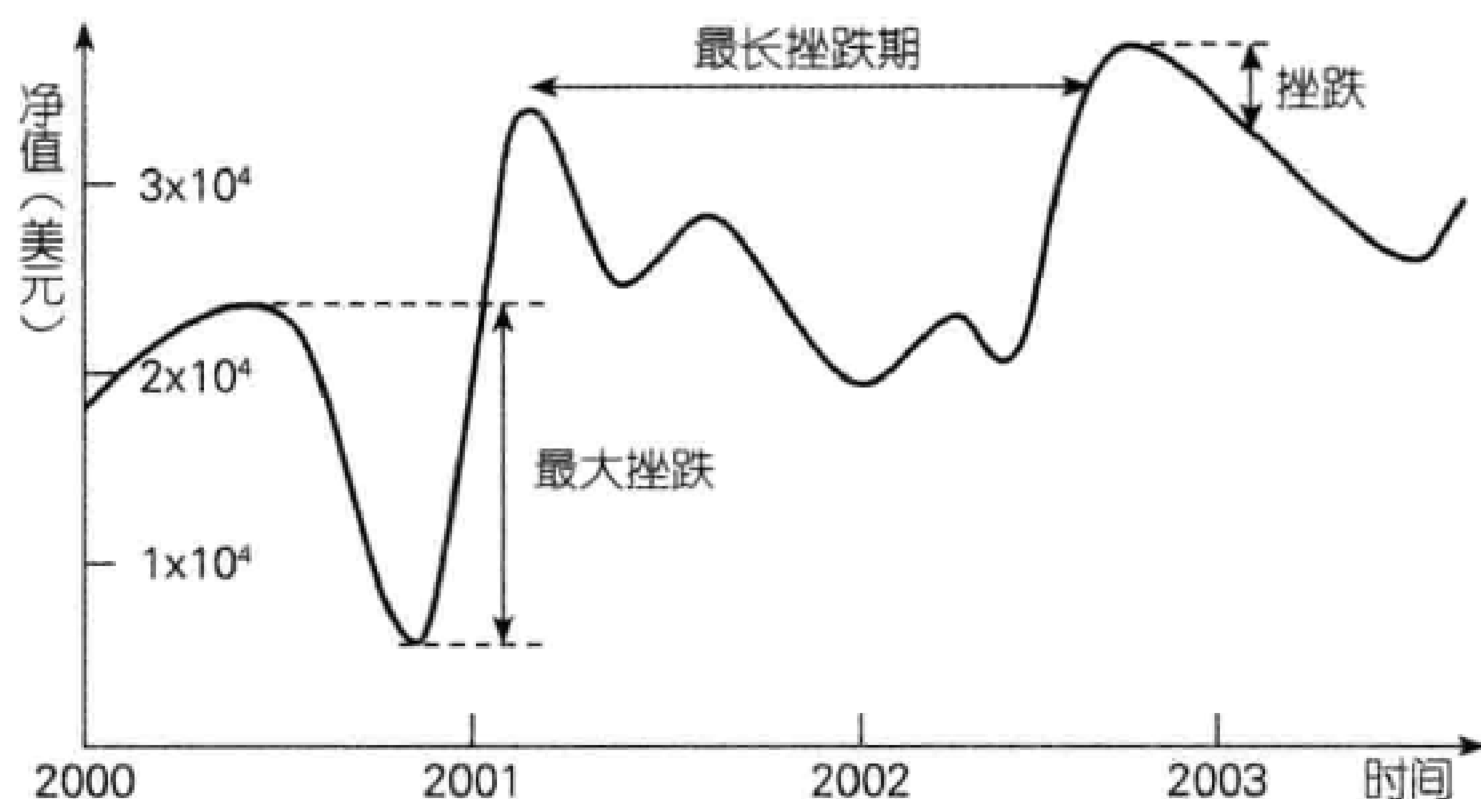


图 2—1　挫跌、最大挫跌和最长挫跌期

根据经验规则：任何夏普比率低于 1 的策略都不适合单独使用。几乎每月都实现盈利的策略，其（年化）夏普比率通常大于 2；几乎每天盈利的策略，其夏普比率通常大于 3。下一章的例 3.4、例 3.6 和例 3.7 中将会详细介绍如何计算不同策略的夏普比率。

2.2.2　挫跌多深、多久?

如果一项策略近期正在亏钱，它就正在经历挫跌。时刻 t 的挫跌被定义为：当前净值（假定期间内未发生任何赎回或注资）与 t 时刻或之前的净值曲线最大值之差。“最大挫跌”是指净值曲线最大值与之后的净值曲线最小值（时间顺序很重要：最小值必须是在最大值之后）之差。净值的最大值又被称为“高水位线”。“最长挫跌期”是指净值重返亏损前的水平所花费的最长时间。

挫跌通常用百分比表示，分母是高水位线时的净值，分子是高水位线

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

以来的净值亏损。

图 2—1 显示了净值曲线的典型挫跌、最大挫跌和最长挫跌期。我在例 3.5 中会详细讲解如何根据每日盈亏表，运用 Excel 或 MATLAB 计算这些数值。记住一件事：最大挫跌和最长挫跌期通常并不发生在同一时间段里。

挫跌的数学定义抽象生僻。然而在现实中，对一个交易员来说，没有比遭遇挫跌更令人痛心疾首、更真实的了。（无论是独立交易员还是机构交易员都如此。当一个机构交易团队遭受挫跌的时候，所有成员都会觉得生活失去了意义，他们终日都在为策略可能会就此失效而忧心忡忡。）所以，挫跌是所有交易员都想尽力最小化的。你必须认真地问问自己，在投资组合清盘或策略结束前，你能承受多深和多久的挫跌？是 20%和 3 个月，还是 10%和 1 个月？将你能够承受的底线与备选策略回测的结果进行比较，从而选择最适合你的那个策略。

即使策略的作者没有公布挫跌的准确数字，你也应该能从净值曲线图上做出大致估计。例如，在图 2—1 中，很容易看出最长挫跌期从 2001 年 1 月至 2002 年 8 月。所以，最长挫跌期大约为 20 个月。同样，最大挫跌开始时净值为 23 000 美元，结束时净值为 5 000 美元。所以最大挫跌为 18 000 美元。

2.2.3　交易成本如何影响策略?

每次基于策略的证券买卖都会产生交易成本。交易越频繁，交易成本对策略的盈利影响就越大。交易成本不仅包括经纪商收取的佣金，还包括流动性成本——当你以市场价格买卖证券时，要支付买卖差价。如果你采用限价指令买卖证券，确实可以避免流动性成本，但却要承担机会成本。因为限价指令可能不被执行，并因此错失了交易的潜在盈利。同样，如果你买卖大笔证券，不可能在不影响价格的情况下完成整个交易。（有时，仅仅是出现大额的买单，哪怕尚未实际买入一股股票之前，就能拉高股票

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

价格！）这种由自身指令所引起的市场价格变动效应被称为“市场冲击”。当证券流动性不好时，它往往会是交易成本的大头。

最后，由于互联网或各种软件的滞后，指令通过程序传送给经纪商的时刻，与指令在交易所被执行的时刻之间存在滞后。这种滞后会造成“滑价”，即触发指令的价格和执行价格之间的差。当然，滑价有正有负，但平均来说，对交易员是成本而不是收益。（如果你觉得滑价是收益，我建议你修改程序，在发送指令的时候，故意延迟几秒钟！）

不同证券的交易成本相差也很大。如果你的指令规模不比最佳买卖报价的平均规模大很多，交易成本通常可以用平均买卖差价的一半再加上佣金来估计。例如，标准普尔 500 指数成分股的平均交易成本（不考虑佣金，这取决于你的经纪商）大约是 5 个基点（即万分之五）。这里，对一个交易循环中的买入和卖出计作两次交易成本。一个先买后卖的来回交易被视作两笔交易，因此，这个例子中的一个来回交易的成本是 10 个基点。E-迷你标准普尔 500 期货合约的交易成本大约是 1 个基点。有时，策略的作者会说明他们的回测业绩中考虑了交易成本，但更多时候并不会这么做。如果没有说明，你就得假设回测结果没有考虑交易成本，从而得判断在考虑成本后策略是否仍然有效。

交易成本影响策略的一个例子是，E-迷你标准普尔 500 期货合约的简单均值回归策略。这项策略基于布林线指标：每当价格超过移动平均值的正负两个标准差时，相应卖空或者买入。当价格回到移动平均值的正负一个标准差之内时清仓。如果允许每 5 分钟建仓和清仓一次，你会发现不计交易成本时夏普比率大约为 3——当然非常出色！不幸的是，如果扣除 1 个基点的交易成本，夏普比率则变成了 -3，变成了完全无利可图的策略。

交易成本影响策略的另一个例子可见例 3.7。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2.2.4 数据有无存活偏差?

股票价格的历史数据库往往不包括那些由于破产、退市、兼并或收购而消失的股票，因此存在所谓的存活偏差，因为数据库中只有幸存者。(共同基金或对冲基金的历史数据库中也不包含那些已经倒闭的基金，同样存在存活偏差。）使用有存活偏差的数据进行回测是很危险的，因为会夸大策略的历史业绩。这在策略有“价值”偏好时（倾向于买便宜的股票）尤为突出，因为有些股票便宜的原因是公司即将破产。如果你的策略只包含那些当时非常便宜但却存活下来的股票（而且可能已经兴盛），而忽略那些退市的股票，回测的结果当然会比交易员当时的实际经历要好得多。

所以，当你读到一个有着良好业绩的“便宜时买进”的策略时，有必要问问策略的作者，回测是否使用的是无存活偏差的数据。否则的话，要谨慎对待（见例 3.3 的策略)。

2.2.5 策略的业绩如何随时间变化而变化?

许多策略 10 年前的业绩要远远好于现在，至少从回测结果来看如此。那时，采用量化策略的对冲基金并不多。并且，买卖差价也比现在大得多，所以，如果用今天的交易成本做回测，会不现实地抬高早年的收益。

有存活偏差的数据也会使得早年业绩过好。存活偏差影响早年业绩的原因在于，回测回溯的越早，消失的股票也就越多。有些股票消失是因为已经破产倒闭，一个纯多头策略在回测的早期会比当年实际的盈亏看起来好。因此，在判断一项策略的适用性时，要重点关注其最近几年的业绩，而不要被包括早年光鲜数字的总体业绩欺骗。

最后，金融市场的“状态转换”，也意味着早年的金融数据并不能简单地应用于今天的相同模型。证券市场监管的变化（如第 5 章中提到的股票价格采用十进制报价或卖空规定的取消）或其他宏观经济事件（如次贷

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

市场的崩溃），都能导致重要的制度转换。

很多具备统计思维的读者很难接受这一点。他们中的很多人会认为，数据越多，回测在统计上就越可靠。但这其实只对由平稳过程产生的金融时间序列来说才是正确的。不幸的是，由于前面提到的原因，金融时间序列显然是非平稳的。

将状态转换纳入一个复杂的“超级”模型是可能的（例 7.1 中我会讨论这个问题），但如果只要求模型相对近期的数据业绩良好，会使事情简单很多。

2.2.6 策略是否存在数据迁就偏差？

如果你构建一个有 100 个参数的策略，完全可能通过优化参数，使历史业绩看起来非常棒。同样可能的是，该策略的未来业绩与回测结果截然不同、非常糟糕。这么多参数，只是使模型与过去发生但未来不会再重现的任何偶然历史事件吻合。实际上，即使只有一两个参数（如建仓和清仓阈值），也很难避免所谓数据迁就偏差，第 3 章会讨论怎样减小数据迁就偏差的影响。一般而言，策略的规则越多，模型的参数越多，就越有可能遭遇数据迁就偏差。能经得起时间考验的往往是简单的模型。

人工智能与选股*

《纽约时报》前些日子发表过一篇文章，介绍了 Ray Kurzweil 先生新创立的一家对冲基金。Ray Kurzweil 先生可称得上是人工智能领域的先驱。（这里要感谢我的博客读者 Yaser Anwar，是他向我推荐了这篇文章。）据 Kurzweil 透露，他的基金的选股决策都是由机器完成的，“……能够观察数十亿笔市场交易，从而发现那些人工无法察觉的模式”（Duhigg，2006）。

虽然我是一名算法交易的信徒，但是当我得知这一交易是基于“人工

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

智能”方法的时候，我对此还是有些怀疑。

我们可以这样简单地理解人工智能（AI），它就是一个尝试使用包含众多参数的函数来拟合历史数据的过程。AI常用的工具有：神经网络、决策树和遗传算法。因为包含的参数很多，我们可以相信它能够捕获到那些人类根本无法察觉到的细微模式。但是，这些模式能够持续吗？这些模式会不会只是一些不会重复的随机噪声？AI领域的专家们向我们保证他们有许多防范措施用以过滤那些瞬间噪声。并且，这些工具也确实在消费者营销和信用卡欺诈检测上效果显著。消费者行为和诈骗行为的模式显然都具有较长的持续期，这使得这些AI算法即使包含大量参数也能有效运行。然而，以我的经验来看，要对金融市场进行预测，这种防范措施是远远不够的，并且对历史数据噪声的过度拟合还会带来严重后果。事实上，我以前也建立过许多基于AI算法的金融预测模型。每次我费尽心力所建立的那些在回测中业绩优异的模型，在之后的实际交易中的业绩却令人十分失望。主要原因可能是，相对于可以获取的大量相互独立的消费者行为和信用交易数据，我们能够获取的在统计学意义上相互独立的金融数据的数量是非常有限的。（你可能会说，我们拥有大量分时金融数据可供使用。但实际上，这些数据是序列相关的，并不是相互独立的。）

但这并不意味着所有基于AI的方法在进行预测时都是无效的。我使用过的有效的AI方法通常具有以下几个特征：

- 基于正确的计量经济学或理论基础，而不是随机发现的模式。
- 所需的参数用到历史数据较少。
- 只用到了线性回归，并未使用复杂的非线性函数。
- 概念上很简单。
- 所有优化都必须在不含未来未知数据的移动回顾窗口中实现，并且这种优化的效果必须不断地被未来未知的数据所证实。

只有满足了这些约束条件的交易模型，我才敢用我那少量且宝贵的历史数据进行检测。显然，“奥卡姆剃刀原理”不仅在科学上有效，在金融

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

上也是如此。

* 改编自我的博客文章《人工智能和选股》，见 epchan. blogspot. com/2006. 12/artificial-intelligence-and-stock. html。

2.2.7 策略是被机构资金管理人忽略吗?

本书的目的在于指导普通投资者从头构建量化交易业务，而并不是指导他们建立一家管理数百万美元的对冲基金，所以，我们并不关心一项策略能否容纳数百万美元。（“容量”是与此相关的一个专业术语，指在对收益率没有负面影响的前提下，一项策略所能容纳的资金量。）实际上恰恰相反，我们应当去寻找那些被大多数机构投资者忽略的策略。例如，由于交易频繁而容量很低的策略，每天只交易极少数股票的策略，持仓时段稀少的策略（如第 7 章中提到的某些商品期货的季节性交易）。这样的特色策略才有利可图，因为它们还没有完全被巨型的对冲基金套利掉。

2.3 小 结

找到潜在的量化交易策略并不难，来自于：

- 商学院或其他经济研究机构的网站。
- 面向零售投资者的金融网站和博客。
- 可与其他交易员交流心得的交易员论坛。

在进行大量的网络搜索及阅读交易杂志之后，你会找到很多有前途的策略。然后，根据我前面所列的筛选标准（更准确的说法应为善意的怀疑），结合自身实际和需求，剔除一些策略：

- 你有多少时间可以用来完善你的交易程序?
- 你的编程能力如何?

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

- 你的资本规模有多大？
- 你的目标是稳定的月度收入还是追求大额的长期资本收益？

在对策略进行详细的回测之前，我们可以先通过一些测试快速淘汰不合适的策略：

- 它能否跑赢基准？
- 它有足够高的夏普比率吗？
- 它有足够小的挫跌和足够短的挫跌期吗？
- 回测有无存活偏差？
- 与早年相比，策略近几年不灵了吗？
- 策略具有避开基金经理激烈竞争的“特色”吗？

完成所有这些快速判断之后，就可以准备进入下一章了。下一章将会对那些选定的策略进行严格的回测，以确保这些策略名副其实。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第3章 回 测

传统投资与量化投资的一个重要区别在于：量化策略可以通过回测来判断其过去的业绩。即便你找到了一个可以提供详尽历史业绩数据的策略，我们还是需要亲自对其进行回测。这样做有几个目的。最基本的，是对策略的重复研究将确保你完整地理解该策略，并在交易系统中对其准确复制以实施。和医学或其他自然科学的研究一样，复制他人的研究结果能有助于我们确认，原始研究中没有犯那些会干扰研究过程的常见错误。但回测不仅仅是要做到应有的谨慎，更重要的是在回测中可以尝试修改原始策略，从而优化并改进策略。

本章将介绍回测的常用平台、回测所用的数据来源、回测必备的业绩度量指标集、应加以避免的常见陷阱，优化和改进策略的简单方法。此外，还会通过几个完整的回测例子来说明所涉及的原理和技术。

3.1 常用的回测平台

有许多商业平台可以用来进行回测，其中有些平台的收费很高。由于

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

本书主要面向的是刚起步的交易员，所以主要介绍一些自己比较熟悉且经济实惠、运用广泛的平台。

3.1.1 Excel

无论对零售交易员还是机构交易员，Excel 都是最基本、最常用的工具。如果你会编写 VB 宏语言，Excel 的功能就更加强大。Excel 的美就在于它能够实现“所见即所得”。由于数据和程序都在同一个界面上使用，所以没有任何隐蔽的内容。另外，在回测中常见的“前视偏差”（稍后会详细介绍）几乎很少在 Excel 中出现（除非使用了宏，这时就无法实现“所见即所得”了），因为我们可以很容易地在电子表格中将不同数据列的日期对齐。Excel 的另一个优势是，能在同一张电子表格中同时进行回测和实时交易，从而避免重复编写程序。Excel 的主要缺点是，它只能对相对简单的模型进行回测。但是，正如上一章所说，简单的模型往往是最好的！

3.1.2 MATLAB

MATLAB（www. mathworks. com）是在大型机构工作的量化研究员和交易员最常用的回测平台之一。它是测试大型股票投资组合策略的理想工具（想象一下在 Excel 中对一个包含 1 500 只股票的策略进行回测——虽说不是不可能，但一定非常痛苦）。MATLAB 集成了许多高级的统计模块和数学模块，如果交易员的交易算法涉及一些复杂而又常用的数学概念，就不用费时费力重新编写程序了。（统计套利交易的因子模型中常用的主成分分析就是一个很好的例子，如果用其他编程语言来实现是非常麻烦的。见例 7.4。）此外，我们还可以免费从网上下载大量 MATLAB 的第三方插件，其中很多插件对量化交易十分有用（例 7.2 中用到的协整软件包就是一个例子）。最后，MATLAB 可以很方便地在网上检索金融信息，并将其转换成有用的格式（称为网络抓取）。例 3.1 展

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

示了这一过程。

尽管 MATLAB 看起来似乎很复杂，实际上还是很容易学的（至少对于基本功能是这样），并且用它很快可以编写一个完整的回测程序。MATLAB 的主要缺点在于它可能相对较贵：一个拷贝大概需要 1 000 美元。不过市场上还有许多 MATLAB 的替代品，编写和使用的代码都与 MATLAB 十分相似，如：

- Matrix (www.omatrix.com)
- Octave (www.gnu.org/software/octave)
- Scilab (www.scilab.org)

这些替代软件的价格只有几百美元，甚至是免费的。当然，收费越高的替代软件，对 MATLAB 编写的程序的兼容性也越好。（如果是完全自己编写程序，不用第三方提供的代码，也就不存在兼容性的问题了。但这样的话，就丧失了使用这门程序语言的一个主要优势。）MATLAB 另一个缺点是，虽然适合进行回测，但却很不适合作为执行平台。所以，在完成策略回测后，还需要用其他程序语言构建一个独立的执行系统。虽然 MATLAB 有这些缺点，但这并不妨碍它在量化交易领域的广泛应用。在本书中，我会给出所有回测的例子所用到的 MATLAB 代码，并在附录中对 MATLAB 语言做简要介绍。

例 3.1　用 MATLAB 从网页中抓取金融数据

MATLAB 不仅能用于数学计算，同时也能用于文本解析。下面是用 MATLAB 从雅虎财经检索股票历史价格信息的例子：

```
clear; % 删除之前定义的变量
symbol='IBM'; % 股票代码
% 检索网页
historicalPriceFile=...
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
urlread(['http://finance.yahoo.com/q/hp? s=', symbol]);
% 将日期存入单元格数组
dateField=...
regexp(historicalPriceFile, ...
'<td class="yfnc_tabledata1" nowrap align=" right" >...
( [\d\w-] +) </td>','tokens');
% 将数值存入单元格数组
numField=regexp (historicalPriceFile, ...
'<td class=" yfnc tabledata1" align=" right" >...
( [\d\.,] +) </td>','tokens');
% 转换为字符串单元格数组
dates= [dateField {:}]';
numField= [numField {:}]';
% 转换为双精度浮点单元格数组
op=str2double (numField (1: 6: end)); % 开盘价
hi=str2double (numField (2: 6: end)); % 最高价
lo=str2double (numField (3: 6: end)); % 最低价
cl=str2double (numField (4: 6: end)); % 收盘价
vol=str2double (numField (5: 6: end)); % 成交量
adjCl=str2double (numField (6: 6: end)); % 调整收盘价
```

这个程序文件可以从 epchan.com/book/example3_1.m 下载，用户名和密码都是“sharperatio”。这个网页抓取程序有一个局限：每次只能抓取一个网页上的信息。由于雅虎财经上的历史数据显示在好多页上，所以使用 MATLAB 检索 IBM 全部历史价格数据并不十分有用。当然，这个程序只是 MATLAB 文本解析功能的简单展示。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

3.1.3 TradeStation

TradeStation（www.tradestation.com）是许多零售交易员很熟悉的经纪商，提供直接与经纪商的服务器相连的、既可回测又可执行交易的一站式平台。这个平台的主要优势有：

- 平台自身就有进行回测所需的几乎全部历史数据，如果使用 Excel 或 MATLAB，要从其他地方单独下载这些数据。
- 在对程序进行回测后，就能立即使用同一程序下单，并将指令传送给经纪商。

这个平台的不足之处在于，一旦用这个软件为策略编写了程序，就自动默认 TradeStation 为经纪商。另外，TradeStation 的专有语言并不能被 MATLAB 这样的软件所兼容，也不包含那些交易员经常使用的统计或数学模块。不过，如果你喜欢一个能够集所有功能于一身的系统，TradeStation 也许是一个不错的选择。

因为我在实际工作中没有用过 TradeStation，所以就不给出用 TradeStation 进行回测的例子了。

3.1.4 高端回测平台

如果你有足够的财力，可以购买某些机构所使用的高端回测平台，下面是部分清单：

- FactSet's Alpha Testing（www.factset.com/products/directions/qim/alphatesting）
- Clarifi's ModelStation（www.clarifi.com/ModelStation-Overview.php）
- Quantitative Analytics' MarketQA（www.qaisoftware.com）
- Barra's Aegis System（www.mscibarra.com/products/analytics/aegis）
- Logical Information Machines（www.lim.com）

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

• Alphacet's Discovery (www.alphacet.com)

这些平台，我只用过 Logical Information Machines 和 Alphacet Discovery。以我 10 年前的经验，Logical Information Machines 测试期货交易策略最合适不过，但测试股票交易策略就要差一些。Alphacet Discovery 是一个集合了数据检索、回测、基于机器学习算法的优化以及自动执行功能的新产品。它的功能非常强大，可以在期货、股票、外汇等市场上进行回测和交易。例 7.1 就是一个使用 Discovery 平台回测的例子。

3.2 查找与使用历史数据库

如果你的头脑中有一个需要特定种类历史数据的策略，首先要做的就是用 Google 去搜索这类数据。你会惊讶地发现，可以找到不少免费的或非常便宜的各类历史数据库（例如，尝试搜索关键词“免费的期货日内价格历史数据”）。表 3—1 列出了这些年来我找到的比较有用的数据库网站，绝大多数都是免费的或非常便宜。我删掉了那些价格昂贵的数据库，如彭博、道琼斯、FactSet、汤森路透和 Tick Data。尽管它们几乎拥有你想要购买的所有数据，但这些数据大都是提供给有实力的机构的，不在个人或新设机构的价格考虑范围。

表 3—1　**用于回测的历史数据库**

来源	优点	缺点
	股票日数据	
Finance.yahoo.com	免费。分拆及股息调整	有存活偏差。一次只能下载一只股票
HQuotes.com	便宜。与 Finance.yahoo.com 数据相同。可下载多只股票	有存活偏差。只有分拆调整，无股息调整
CSIdata.com	便宜。Yahoo! 和 Google 的源数据。可下载多只股票	有存活偏差

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

续表

来源	优点	缺点
TrackData. com	便宜。分拆及股息调整。可下载多只股票。提供基本面数据	有存活偏差
CRSP. com	无存活偏差	昂贵。每月更新一次
	期货日数据	
Quotes-plus. com	便宜。可下载多只期货	
CSIdata. com	同上	
	外汇日数据	
Oanda. com	免费	
	股票日内数据	
HQuotes. com	同上	日内数据的历史不长
	期货日内数据	
DTN. com	NxCore 产品提供买、卖历史数据	昂贵：需订阅实时数据传送
	外汇日内数据	
GainCapital. com	免费。有较长的数据历史	

尽管在互联网上寻找数据来源要比寻找潜在的策略容易得多，但要使用这些数据库还存在很多问题和陷阱，我接下来对此进行详细分析。这些注意事项主要是针对股票和交易型开放式指数基金（ETF），主要包括这么几个方面。

3.2.1 数据是否经分拆及股息调整?

一家公司在除权日 T 分拆其股票，每股分拆为 N 股（N 通常为 2，但也可能是小数，比如 0.5。当 N 小于 1 时，为反向分拆），T 日之前的所有股票价格应该乘上 $1/N$。相应地，当公司在除息日 T 每股派发 d 元的股息，那么 T 日之前的所有股票价格应该乘上（Close（$T-1$）$-d$）/Close（$T-1$），其中 Close（$T-1$）是 T 日前一交易日的收盘价格。注意，调整股票历史价格时，是乘上一个因子而不是减去 d 元，以使调整

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

前后股票日收益率不变。这是最常见的一种方式，雅虎财经就是这么做的。(如果减去 d 元，虽可保证调整前后股票日价格变化相同，但日收益率会变化。）如果不对历史数据做出调整，除息日的开盘价要比前一日的收盘价低不少（除去正常的市场波动），这很容易导致一个错误的交易信号。

我建议最好找一个分拆及股息调整后的历史数据库，否则就必须找到一个单独包含分拆和股息信息的历史数据库（如 earnings. com），然后自己做调整，而这是非常烦琐且容易出错的工作。见下例。

例 3.2　分拆和股息的调整

现在我们来看 IGE，一只 ETF，历史上既有分拆，也有股息。2005 年 6 月 9 日（除权日），1 股分拆为 2 股。以下是除权日前后没有经过调整的价格（可以从雅虎财经直接下载 IGE 历史价格的 Excel 数据）：

日期	开盘价	最高价	最低价	收盘价
2005—6—10	73.98	74.08	73.31	74
2005—6—9	72.45	73.74	72.23	73.74
2005—6—8	144.13	146.44	143.75	144.48
2005—6—7	145	146.07	144.11	144.11

2005 年 6 月 9 日之前的价格由于分拆而需要调整。由于 N=2，因此将价格乘上 1/2 即可。以下是调整后的价格：

日期	开盘价	最高价	最低价	收盘价
2005—6—10	73.98	74.08	73.31	74
2005—6—9	72.45	73.74	72.23	73.74
2005—6—8	72.065	73.22	71.875	72.24
2005—6—7	72.5	73.035	72.055	72.055

细心的读者会发现，调整后的收盘价与雅虎财经给出的调整后的收盘价还是不一样的。这是因为 2005 年 6 月 9 日之后还有股息，雅虎财经对此也进行了调整。每次股息调整都是乘上一个因子，故总调整的因子就是

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

所有单个因子的乘积。以下是2005年6月9日至2007年11月所有的股息、前一交易日未调整的收盘价及调整因子：

除息日	股息	前一日收盘价	调整因子
2007—9—26	0.177	128.08	0.998618
2007—6—29	0.3	119.44	0.997488
2006—12—21	0.322	102.61	0.996862
2006—9—27	0.258	91.53	0.997181
2006—6—23	0.32	92.2	0.996529
2006—3—27	0.253	94.79	0.997331
2005—12—23	0.236	89.87	0.997374
2005—9—26	0.184	89	0.997933
2005—6—21	0.217	77.9	0.997214

（请在Excel中使用我在上面给出的公式，自行检查调整因子是否满足我在这里给出的数值。）股息总因子为0.998618 ×0.997488 ×…×0.997214 =0.976773，适用于2005年6月9日这一天及之后未调整的价格。分拆和股息的总因子是0.976773 × 0.5 = 0.488386，适用于所有2005年6月9日之前未经调整的价格。以下是用这些因子计算出来的调整后价格：

日期	开盘价	最高价	最低价	收盘价
2005—6—10	72.26163	72.35931	71.6072	72.28117
2005—6—9	70.76717	72.02721	70.55228	72.02721
2005—6—8	70.39111	71.51929	70.20553	70.56205
2005—6—7	70.81601	71.33858	70.38135	70.38135

以下是2007年11月1日雅虎上的调整后数据：

日期	开盘价	最高价	最低价	收盘价	成交量	调整后价格
2005—6—10	73.98	74.08	73.31	74	179 300	72.28
2005—6—9	72.45	73.74	72.23	73.74	853 200	72.03
2005—6—8	144.13	146.44	143.75	144.48	109 600	70.56
2005—6—7	145	146.07	144.11	144.11	58 000	70.38

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

可见，计算得出的调整后收盘价与雅虎上的一致（四舍五入保留小数点后两位）。当然，在你读到本书时，IGE 很有可能又发放了股息，并且又进行了分拆，所以，雅虎上的数据可能已经与上表中的不一样。你可以做做练习，根据新的分拆和股息情况再次进行调整，并与当前雅虎上的数据进行比较。

3.2.2 数据有无存活偏差?

我们在第 2 章中已经讨论过这个问题。不幸的是，对于一个起步者来说，无存活偏差的数据库价格过于昂贵，很难负担得起。能够克服这个困难的一种方法是，自己去收集将来可用于做回测的带时刻标签的数据。如果每天都保存了所有关注的股票的价格，就有了可供将来使用的带时刻标签或无存活偏差的数据库。另一种可以降低存活偏差影响的方法是，只用近期数据对策略进行回测，这样便不会因为有太多消失的股票而影响结果。

例 3.3 存活偏差人为抬高回测业绩的例子

这是一个“买进低价股票”的玩具策略。（警告：这个玩具策略有很大的金融风险！）假设在年初从 1 000 只市值最大的股票中，等值买进价格最低的 10 只股票并持有一年。如果数据库无存活偏差，所得结果如下：

证券代码	2001—1—2 收盘价	2002—1—2 收盘价	最后价格
ETYS	0.2188	无	0.125
MDM	0.3125	0.49	0.49
INTW	0.4063	无	0.11
FDHG	0.5	无	0.33
OGNC	0.6875	无	0.2
MPLX	0.7188	无	0.8
GTS	0.75	无	0.35
BUYX	0.75	无	0.17
PSIX	0.75	无	0.2188

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

除 MDM 外，其他所有股票在这一年间都退市了（毕竟互联网泡沫还是很严重的），故 2002 年 1 月 2 日没有收盘价。最后价格是 2002 年 1 月 2 日或退市前最后成交的价格。这一投资组合的年收益率是—42%。

如果数据库有存活偏差，这个组合就会筛掉这一年间所有退市的股票，进而选中以下股票：

证券代码	2001—1—2 收盘价	2002—1—2 收盘价
MDM	0.3125	0.49
ENGA	0.8438	0.44
NEOF	0.875	27.9
ENP	0.875	0.05
MVL	0.9583	2.5
URBN	1.0156	3.0688
FNV	1.0625	0.81
APT	1.125	0.88
FLIR	1.2813	9.475
RAZF	1.3438	0.25

注意，由于被选中的均是“存活”下来的股票，因此 2002 年 1 月 2 日全都有收盘价。这个投资组合的年收益率是 388%！

此例中，策略的实际收益率是—42%，但由于存活偏差的存在，收益率被虚估为 388%。

3.2.3 策略用最高、最低价吗？

在几乎所有股票的日数据中，最高、最低价的噪声远远大于开盘、收盘价。这意味着，即使你的限价买入指令低于最高价，也可能无法成交，反之，对于限价卖出指令也是如此。（这可能是因为在最高价成交的指令

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

可能很少，或者交易是在一个指令无法送达的市场中发生的，甚至有可能记录下来的最高、最低价是不正确的且没有被筛掉。）所以，使用最高、最低价做回测不如开盘、收盘价可靠。

事实上，有时开盘市价指令或收盘市价指令也无法按数据库中的历史开盘、收盘价成交。因为这些历史价格可能来自大型交易所（如纽约证券交易所），也可能是一个包括所有区域性交易所在内的综合价格。指令送达的地点不同，其成交价格也与数据库中相应的历史开盘、收盘价不同。尽管如此，这种开盘、收盘价的差异对回测业绩的影响，通常比最高、最低价要小，因为后者总是抬高回测收益。

从数据库检索完数据后，应对其进行查错。最简单的方法就是算一下日收益。如果你有最高价、最低价、开盘价、收盘价的数据，还可以算一下不同组合的日收益（如昨天的最高价与今天的收盘价）。对偏离均值 4 个标准差的收益要仔细检查。一般而言，极端收益与消息发布或市场指数异动有关。不然的话，就是数据本身有问题。

3.3 业绩度量

量化交易员有很多不错的业绩度量指标，究竟使用哪些指标，很大程度上取决于个人偏好。但我认为，在策略之间、交易员之间进行横向比较，最重要的两个指标是夏普比率和挫跌。注意，我并没有提到平均年化收益率这个投资者最常用的度量指标，因为如果要使用这个指标，就必须得详细说明计算收益率所用到的分母。例如，在多空策略中，分母用一个方向的资本还是两个方向的呢？收益率是杠杆收益率（分母用账户净值）还是无杠杆（分母用投资组合的市场价值）呢？如果净值或市场价值每日变化，分母是用移动平均值还是每天或每月末的价值呢？使用夏普比率和挫跌，而非标准的业绩度量指标，在对不同策略进行比较时，上述绝大多

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

数（并非全部）问题就能避免。

第 2 章介绍了夏普比率、最大挫跌和最长挫跌期的概念。本章结合 Excel 和 MATLAB 中的例子，指出夏普比率计算中的一些细节问题。

计算夏普比率时，甚至老练的投资经理也会混淆的一个细节问题是：是否需要从货币中性组合的收益中减去无风险收益？答案是否定的。货币中性组合可以用卖空所得现金买入证券，故它是自融资的，其融资成本很小（存贷利率差），在回测时可以忽略。同时，保证金余额能获得与无风险利率 r_F 近似的存款利率。假定策略收益率（组合收益率减去存款利率）为 R，无风险利率为 r_F。所以，用于计算夏普比率的超额收益率就是 $R+r_F-r_F=R$。因此在实际计算中，完全可以忽略无风险利率，只需关注股票头寸的收益率即可。

同样，一个没有隔夜头寸的纯多头当日交易策略也没有资金成本，其超额收益率也是策略本身的收益率。一般而言，在计算夏普比率时，仅当策略需要支付资金成本时，才需要减掉无风险利率。

年化夏普比率可以使策略更加容易。大多数人都知道如何年化平均收益率。例如，年均收益率就是月均收益率的 12 倍。

然而，年化收益率标准差就稍微复杂一点。这里，假设月收益率是序列不相关的（Sharpe，1994），年收益率标准差就是月收益率标准差的 $\sqrt{12}$ 倍。于是，年化夏普比率就是月夏普比率的 $\sqrt{12}$ 倍。

一般而言，假设每个交易时段的长度为 T，T 可以是 1 个月、1 天、1 小时等，若要计算平均收益率、标准差以及相应的年化指标，就必须先算出一年的交易时段数 N_T。然后

年化夏普比率 $=\sqrt{N_T}\times$ 基于 T 的夏普比率

例如，策略只在纽约证券交易所交易时间（美国东部时间 9：30—16：00）持仓，每小时平均收益率为 R，每小时收益率标准差为 s，则年化夏普比率为：$\sqrt{1\ 638}\times R/s$。因为，$N_T=$（252 个交易日）×（6.5 每

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

一交易日交易小时）=1 638。一个常见的错误计算是：$N_T = 252 \times 24 = 6\ 048$。

例 3.4　计算纯多头策略和市场中性策略的夏普比率

先考虑一个关于 IGE 的纯多头策略：2001 年 11 月 26 日以收盘价买入 1 股并持有，2007 年 11 月 14 日以收盘价卖出。假设持有期内无风险利率为每年 4%，从雅虎财经下载持有期内 IGE 的每日价格，保存为 Excel 文件 IGE.xls（不是默认的逗号分隔文件）。使用 Excel 或 MATLAB 的步骤如下：

使用 Excel

1. 下载的数据已在 A—G 列；

2. 按日期升序对所有列进行排序（使用数据排序功能，选择“扩展选定区域”单选按钮，选择“升序”和“有标题行”单选按钮）；

3. 在 H3 单元格内输入“=（G3—G2）/G2”，计算日收益率；

4. 选定 H3 单元格，双击单元格右下角的黑点，使 H 列的整列按 H3 单元格的方法计算 IGE 的日收益率；

5. 在 H1 单元格内输入“日收益率”以便阅读；

6. 在 I3 单元格内输入“=H3—0.04/252”，计算超额日收益率，假设无风险利率为每年 4%，一年有 252 个交易日；

7. 选定 I3 单元格，双击单元格右下角的黑点，使 I 列的整列按 I3 单元格的方法计算超额日收益率；

8. 在 I1 单元格内输入“超额日收益率”以便阅读；

9. 在 I1506 单元格内输入“=SQRT（252）* AVERAGE（I3：I1505）/STDEV（I3：I1505）”；

10. I1505 单元格显示结果应该是“0.789317538”，这个结果是购买并持有策略的夏普比率。

在我的个人网站上有已经做好的电子表格 epchan.com/book/

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

example3 4. xls。

使用 MATLAB

```
%清除工作空间已有变量
clear;
%从"IGE.xls"读入数据
[num, txt]=xlsread('IGE');
%第一列(从第二行开始)是交易日,格式为 mm/dd/yyyy.
tday=txt(2:end, 1);
%将时间格式转化为 yyyymmdd.
tday=datestr(datenum(tday,'mm/dd/yyyy'),'yyyymmdd');
%将数据字符串转化为单元型数组,再转化为数值型变量
tday=str2double(cellstr(tday));
%最后一列是调整后的收盘价
cls=num(:, end);
%将 tday 按时间升序排列
[tday sortIndex]=sort(tday,'ascend');
%将 cls 按时间升序排列
cls=cls(sortIndex);
%计算日收益率
dailyret=(cls(2:end)-cls(1:end-1))./cls(1:end-1);
%假设无风险利率为 4%/年,每年 252 个交易日,计算超额收益率;
excessRet=dailyret-0.04/252;
%输出结果应该是 0.7893
sharpeRatio=sqrt(252)*mean(excessRet)/std(excessRet)
```

MATLAB 代码可以从我的个人网站上下载 epchan.com/book/example3 4. m。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

再考虑多空市场中性策略。这个策略由前面策略变化而来：在买入IGE时，做空等量美元的标准普尔存托凭证（SPY）进行对冲，2007年11月14日同时将这两个头寸平仓。从雅虎财经下载SPY价格数据，保存为SPY.xls文件。Excel和MATLAB的使用步骤与上面很相似，留给读者做练习写出其精确步骤：

使用Excel

1. 按上面的方法对SPY.xls的列按日期升序排序；
2. 复制SPY.xls的G列（调整后收盘价），粘贴到IGE.xls的J列；
3. 检查J列的行数是否与A-I列一样。如果不一样，说明数据有问题，核实从雅虎财经下载的数据是否是同一日期范围内的数据；
4. 使用同样的方法在K列计算日收益率；
5. 在K列的首行输入“日收益率”以便阅读；
6. 在L列，计算H和K的差并除以2，得到套期保值策略的日净收益率（除以2是因为使用了2倍的资本）；
7. 在L1506单元格内计算对冲策略的夏普比率，结果应该是“0.783681”。

使用MATLAB

```
% 接上例MATLAB代码
% 在此插入你自己的代码，像上例那样从SPY.xls读入数据
% 将包含SPY每日收益的数组命名为“SPY日收益率”
% 日净收益率（除以2是因为使用了2倍的资金）
netRet=（dailyret-dailyretSPY）/2;
% 输出结果应该是0.7837.
sharpeRatio=sqrt(252)*mean(excessRet)/std(excessRet)
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例 3.5 计算最大挫跌和最长挫跌期

继续使用上面的多空市场中性策略，来说明如何计算最大挫跌和最长挫跌期。第一步需要在每日收盘时计算“高水位线”，即到这一天为止，策略的最大累积收益率。（用累积收益率曲线来计算高水位线和挫跌，与用净值曲线是一样的，因为净值正是初始投资额与“1＋累积收益率”的乘积。）有了高水位线，我们就可以计算出挫跌、最大挫跌和最长挫跌期。

使用 Excel

1. 在 M3 单元格内输入“＝L3”；

2. 在 M4 单元格内输入“＝（1＋M3）*（1＋L4）－1”，计算策略到某天为止的累积收益率。按此方法计算 M 列，得到策略的累积复合收益率，并删除该列最后一个单元格，将这列命名为“累积收益率”；

3. 在 N3 单元格内输入“＝M3”；

4. 在 N4 单元格内输入“＝MAX（N3，N4）”，计算到某天为止的“高水位线”。按此方法计算 N 列并删除该列最后一个单元格，将这列命名为“高水位线”；

5. 在 O3 单元格内输入“＝（1＋N3）/（1＋M3）－1”，计算每日收盘后的挫跌，按此方法计算 O 列，得到策略的挫跌；

6. 在 O1506 单元格内输入“＝MAX（O3：O1505）”，计算策略的最大挫跌，结果约为 0.1053，即最大挫跌是 10.53%；

7. 在 P3 单元格内输入“＝IF（O3＝0，0，P2＋1）”，计算当前挫跌期。按此方法计算 P 列，是为策略的挫跌期，并删除该列最后一个单元格；

8. 在 P1506 单元格内输入“＝MAX（P3：P1505）”，计算策略的最长挫跌期，这个值大概是 497，即最长挫跌期含 497 个交易日。

使用 MATLAB

% 接上例 MATLAB 代码

% 累积收益率

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
cumret=cumprod(1+netRet)-1;plot(cumret);
[maxDrawdown maxDrawdownDuration]=...
calculateMaxDD(cumret);
[maxDrawdown maxDrawdownDuration]=...
calculateMaxDD(cumret);
% 最大挫跌，输出结果应该是 0.1053
maxDrawdown
% 最长挫跌期，输出结果应该为 497
maxDrawdownDuration
```

注意，上述代码调用了子程序“calculateMaxDrawdown”，内容所示：

```
function [maxDD maxDDD]=calculateMaxDD(cumret)
% [maxDD maxDDD]=calculateMaxDD(cumret)
% 在累积收益率的基础上计算最大挫跌和最长挫跌期
% 将高水位线初始化为 0.
highwatermark=zeros(size(cumret));
% 将挫跌初始化为 0.
drawdown=zeros(size(cumret));
% 将最长挫跌期初始化为 0.
drawdownduration=zeros(size(cumret));
for t=2:length(cumret)
 highwatermark(t)=max(highwatermark(t-1), cumret(t));
 % 计算每日挫跌(相对于高水位)
 drawdown(t)=(1+highwatermark(t))/(1+cumret(t))-1;
   if (drawdown(t)==0)
      drawdownduration(t)=0;
   else
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
        drawdownduration(t)=drawdownduration(t-1)+1;
    end
end
maxDD=max(drawdown); % 最大挫跌
maxDDD=max(drawdownduration); % 最长挫跌期
```

这个程序文件可在 epchan.com/book/calculateMaxDD.m 下载。图 3—1中的这段累积收益率曲线显示了最大挫跌和最长挫跌期。

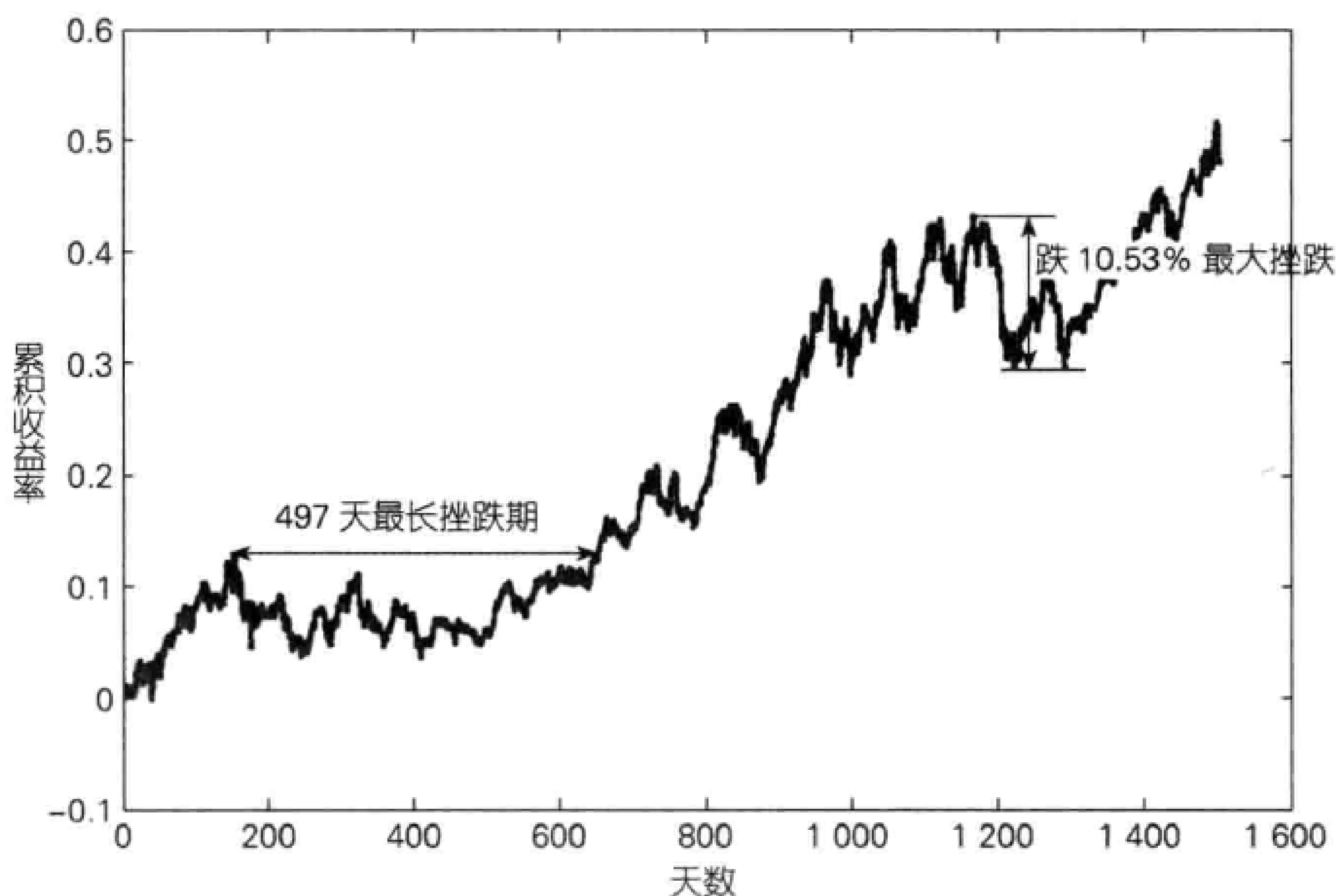


图 3—1 例 3.4 的最大挫跌和最长挫跌期

3.4 避免常见的回测陷阱

基于过去发生的历史信息产生历史交易，并考察这些交易业绩的过

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

程，称为回测。计算机算法产生交易的回测看似简单，但实际上很容易出错。历史业绩高估（相对于已发生的现实交易）是回测的常见错误。我们已经知道，使用有存活偏差数据会导致回测业绩高估。然而，还有一些与如何编写回测程序以及如何构造交易策略相关的常见回测陷阱。我介绍其中最常见的两种，以及规避它们的方法。

3.4.1 前视偏差

前视偏差是指使用交易完成之后的信息。例如，“在日最低价的 1% 之内买入股票”的交易规则，就有前视偏差，因为在当日市场收盘前，是不可能知道日最低价的。又如，使用全部数据回归得来的系数，来产生一个基于前后两段价格序列的线性回归模型的交易信号，同样有前视偏差。

我们如何避免前视偏差呢？使用“滞后”的历史数据来计算策略信号，可以避免前视偏差。滞后数据系列意味着，在计算移动平均值、最高价、最低价、成交量等指标时，只使用“上一”交易期限的收盘数据。(当然，如果策略只在交易期限结束时触发，那就无需使用滞后数据。)

使用 Excel 等“所见即所得”的程序比 MATLAB 更容易避免前视偏差。因为在 Excel 中很容易将不同列的数据对齐，并确保每个单元格里的公式只使用当前行之前的数据。Excel 的单元格高亮显示功能，可以让我们直接查看使用当日数据所产生交易信号的情况。(双击带公式的单元格，该公式所使用数据所在的单元格会高亮显示。)

即便非常谨慎，在编写回测程序时还是有可能犯前视偏差错误。有些错误非常微妙、难以避免，特别是在使用 MATLAB 的时候。最好用下列方法对回测程序做最后检查：使用所有历史数据运行程序，把推荐头寸存入文件 A（目标文件包含程序每天生成的所有推荐头寸）。然后移除最近 N 天的历史数据，即如果原始数据中的最后一天是 T 日，移除后的数据中的最后一天就是 $T-N$ 日，N 可以是 10 天到 100 天不等。再次运行回测程序并将结果存入另一个文件 B。移除文件 A 的最后 N 行，此时文件

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

A 与文件 B 的行数（天数）相同，最后一天均为 $T-N$ 日。最后，比较文件 A 和文件 B 中的头寸。如果头寸不一致，说明回测程序中存在前视偏差，必须找出并改正，因为头寸不一致意味着，误将那 N 天移除了的数据也加入了文件 A 中头寸的计算。我将在例 3.6 中介绍这个稍微有点绕的过程。

3.4.2 数据迁就偏差

第 2 章提到，因迁就历史数据的噪声而过度优化模型参数，造成策略的回测业绩高于未来业绩，即为数据迁就偏差。在基于历史数据的统计预测模型中，数据迁就偏差非常普遍，而在金融中，因为独立数据的量非常有限，所以这个问题就更加严重。高频数据倒是非常多，但只对高频模型有用。尽管股票市场数据可以追溯到 20 世纪初，但只有最近十年的数据才适用于构建预测模型。此外，如果发生了第 2 章讨论过的状态转换，最近几年的旧数据可能也没法用。独立数据的量越少，交易模型中用到的可调整参数就应该越少。

根据经验，我在计算移动平均值时不会超过 5 个参数，包括如建仓清仓阀值、持有期限、回溯期限等定量指标。此外，并非所有的数据迁就偏差都是由参数优化造成的。用相同数据集重复做回测时，还会影响模型的定性决策选择，如在开盘还是收盘时建仓、是否隔夜持仓、交易大盘股还是中盘股等。通常，这些可以优化回测业绩的定性选择，对未来未必是最优的。构建数据驱动模型时，几乎不可能完全消除数据迁就偏差，但有几种方法可以降低这种偏差。

样本含量

降低数据迁就偏差的最基本方法是，按照要优化的自由参数个数，使用足够多的回测数据。根据经验规则，通常假定优化参数所需的数据点个数，是模型中自由参数个数的 252 倍（一年交易天数为 152 天。这个假定未基于任何统计调查结果，完全是由经验所得）。例如，回测三参数的当

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

日交易模型，至少要用三年的日价格数据。如果是分钟交易模型，则至少需要7个月（252/390年）的分钟数据（每天交易分钟数6.5×60=390）。注意，如果是一个日交易模型并且有了7个月的分钟数据点，实际上也只有7×21=147个有效数据（每月交易天数21），要回测三参数模型是不够的。

样本外测试

将历史数据根据时间先后分为两段，后一段数据用于样本外测试。构建模型时，参数优化和定性选择使用前一段数据（称为“训练集”），所得模型的测试使用后一段数据（称为“测试集”）。（两段数据的大小要大致相等，若数据量不够，则至少用三分之一的数据做测试。训练集的最小含量由上一节提到的经验规则决定。）在理想情况下，基于训练集的最优参数和决策，对于测试集也是最优的，不过实际上很难做到这一点。但测试集上的业绩起码要合理。否则，模型就存在数据迁就偏差，需要进一步简化并减少参数。

动态参数优化是更严密、也更复杂的样本外测试方法。这种方法通过使参数不断适应变化的历史数据，来消除数据迁就偏差。（参见“无参数交易模型”）

无参数交易模型*

我曾为之工作过的一个投资经理，很热衷于宣传他的交易模型“没有任何自由参数”。由于业内的保密行规，他没有透露过更多的技术细节。

后来，我开始明白无自由参数交易模型的含义了，并不是说不包含任何参数，如不包含计算价格趋势时要用的回溯期或确定建仓或清仓时间的阈值，这是不可能的。真正的含义是所有这些参数都在一个移动回溯窗口中进行动态优化。假如你问：“模型中有固定的盈利上限吗?”交易员会诚实地回答：“没有，盈利上限并不是一个输入参数，它是由模型本身所决定的。”

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

无参数交易模型的优势在于，能够减少多参数输入模型的过度优化风险（数据迁就偏差）。因此，回测业绩应与未来交易的实际业绩更加吻合。

（注意，参数优化并不一定是挑选那个回测业绩最好的参数集。通常，更应该是基于不同参数集的某种平均来做交易决定。）

在下达下一条指令前，及时对所有参数进行优化，是对计算的一个很大考验，若要进行多维优化就更加困难了。因此，在开始研究例 7.1 所述的状态转换模型之前，我自己极少使用无参数交易模型。状态转换模型基本上是无参数的（我在优化时忽略了一些参数，不是因为任何技术上的困难，而是时间有限）。在这个例子中，我怎么在几分钟之内完成回测参数优化的呢？我使用了高端回测交易平台（Alphacet Discovery）。

* 改编自我的博客文章《无参数交易模型》，见 epchan. blogspot. com/2008/05/parameterless-trading-models. html。

最靠谱的样本外测试方法，很多交易员都很熟悉，叫做“仿真交易”。用尚未发生的真实数据运行模型，是最可靠的测试方式。仿真交易不仅能做真实准确的样本外测试，同时也常常可以发现模型中的前视偏差，以及各种与操作相关的问题。第 5 章将详细讨论仿真交易。

对于一个需要通过测试来核实的公开策略，从策略的公开日到测试日的这段时间，是不折不扣的样本外测试期。只要不在样本外测试期优化模型的参数，这段时间跟策略的仿真交易的效果一样好。

例 3.6　GLD 和 GDX 的配对交易

本例展示如何将数据分成训练集和测试集。我们将回测一个配对交易策略，在训练集上优化参数，在测试集上观察效果。

GLD 代表黄金的现货价格，GDX 是一揽子采金企业股票，两者的价格是高度相关的，故 GLD 和 GDX 可用于做配对交易。我在博客中运用协整分析对 ETF 的这一配对交易进行了详细讨论（见

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

epchan. blogspot. com/2006/11/reader-suggested-possible-trading. html ）。不过我要到第 7 章才会讲训练集上的协整分析，结果表明，GLD 多头和 GDX 空头所形成的差价呈均值回归。通过训练集上的回归分析可得出 GLD 和 GDX 之间的对冲比率，并设定配对交易策略进出市场的阀值。从后面可以看到，阀值在训练集上的优化会改变策略在测试集上的业绩。(程序文件可从 epchan. com/book/example3_ 6. m 下载，数据文件为 GLD. xls 和 GLD. xls。) 程序使用滞后命令，将时间序列滞后一期，这在 epchan. com/book 中也能找到。还会使用“普通最小二乘法 (OLS)”命令进行线性回归，在 spatial-econometrics. com 上可免费打包下载。

使用 MATLAB

```
% 清除工作空间已有变量
clear;
% 将"GLD. xls" 读入 MATLAB
[num, txt]=xlsread('GLD');
%第一列(从第二行开始)是交易日,格式为 mm/dd/yyyy
tday1=txt(2:end, 1);
% 将时间格式转化为 yyyymmdd
tday1=...
datestr(datenum(tday1, 'mm/dd/yyyy'), 'yyyymmdd');
% 将数据字符串转化为单元型变量,再转化为数值型变量
tday1=str2double(cellstr(tday1));
% 最后一列是调整后收盘价
adjcls1=num(:, end);
% 读入"GDX. xls"
[num, txt]=xlsread('GDX');
%第一列(从第二行开始)是交易日,格式为 mm/dd/yyyy
tday2=txt(2:end, 1);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 将时间格式转化为 yyyymmdd
tday2=...
datestr(datenum(tday2, 'mm/dd/yyyy'), 'yyyymmdd');
% 将数据字符串转化为单元型变量,再转化为数值型变量
tday2=str2double(cellstr(tday2));
% 最后一列是调整后收盘价
adjcls2=num(:, end);
% 找到两组数据的交集并按升序排列
[tday, idx1, idx2]=intersect(tday1, tday2);
cl1=adjcls1(idx1);
cl2=adjcls2(idx2);
trainset=1:252; % 定义训练集下标
% 定义测评集下标
testset=trainset(end)+1:length(tday);
% 用回归函数计算得到对冲比率
results=ols(cl1(trainset), cl2(trainset));
hedgeRatio=results. beta;
% 差价= GLD-对冲比率 * GDX
spread=cl1-hedgeRatio * cl2;
plot(spread(trainset));
figure;
plot(spread(testset));
figure;
% 训练集平均差价
spreadMean=mean(spread(trainset));
% 训练集差价标准差
spreadStd=std(spread(trainset));
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 差价标准化(用 z-scores 方法)
zscore=(spread-spreadMean)./spreadStd;
% 在组合价值向下跌破 2 倍标准差时,购买此差价组合
longs=zscore<=-2;
% 当组合价值上升超过 2 倍标准差时,做空该差价组合
shorts=zscore>=2;
% 当组合价值回到 1 倍标准差以内时,清仓
exits=abs(zscore)<=1;
% 初始化头寸数组
positions=NaN(length(tday), 2);
% 多头入市
positions(shorts, :)=repmat([-1 1], [length(find(shorts)) 1]);
% 空头入市
positions(longs, :)=repmat([1 -1],[length(find(longs)) 1]);
% 清仓
positions(exits, :)=zeros(length(find(exits)), 2);
% 确保继续持仓,除非出现清仓信号
% positions=fillMissingData(positions);
cl=[cl1 cl2]; % 合并两个价格序列
dailyret=(cl - lag1(cl))./lag1(cl);
pnl=sum(lag1(positions).*dailyret, 2);
% 训练集的夏普比应该是 2.3
sharpeTrainset=...
sqrt(252)*mean(pnl(trainset(2:end)))./std(pnl(trainset(2:end)))
% 测评集的夏普比应该是 1.5
sharpeTestset=sqrt(252)*mean(pnl(testset))./std(pnl(testset))
plot(cumsum(pnl(testset)));
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 保存头寸文件以便检查数据先窥偏差
save example3_6_positions positions;
```

文件 lag1.m：

```
function y=lag1(x)
% y=lag(x)
if (isnumeric(x))
% 第一个元素填充为 NaN
y=[NaN(1,size(x,2));x(1:end-1, :)];elseif (ischar(x))
% 第一个元素填充为"
y=[repmat("",[1 size(x,2)]);x(1:end-1, :)];else error('Can only be
numeric or char array');
End
```

此配对交易策略在训练集和测试集上的夏普比率都很高，因此可认为它是无数据迁就偏差的。但也许还有进一步改进的空间。若把建仓阈值改为 1 倍标准差、清仓阈值改为 0.5 倍标准差，训练集上的夏普比率会上升到 2.9，测试集上的夏普比率会上升到 2.1。显然，这一阈值集更佳。

不过，在训练集上进行参数优化也许会降低测试集上的业绩。这种情况下，应选择使得训练集和测试集上的业绩结果都较好（也许不是最好）的参数集。

我没有将交易成本考虑在内（下一节会讨论交易成本）。读者可以自己做练习。由于这一策略不是频繁交易，因此，交易成本对所得的夏普比率影响并不大。

为观察这一策略是如何工作的，读者可参见图 7.4 所显示的差价，第 7 章将结合平稳性和协整性进行进一步讨论。你将看到差价走势呈现出很明显的均值回归。因此，不断地低买高卖是很管用的。

最后，还要检测任何可能的数据前视偏差。在上面的 MATLAB 代码“cl2=adjcls2 (idx2);”之后，添加以下代码。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 将最近的交易日数据移除掉
cutoff=60;% 移除最近 60 天的数据
tday(end-cutoff+1:end, :)=[];
cl1(end-cutoff+1:end, :)=[];
cl2(end-cutoff+1:end, :)=[];
```

将以下代码放在上面 MATLAB 程序的结尾，取代“save example3_6_positions positions”。

```
% 检测数据先窥偏差的第二步
oldoutput=load('example3_6_positions');
oldoutput.positions(end-cutoff+1:end, :)=[];
if (any(positions~=oldoutput.positions))
fprintf(1, 'Program has look-forward-bias! \n');
End
```

将新的代码保存为“example3_6_1.m”并运行，你会发现“Program has look-forward-bias”语句不会被打印出来，这说明算法通过测试了。

敏感性分析

在完成模型的参数和各种特征的优化、通过测试集的检验之后，可以通过改变这些参数或改变模型的定性决策，来观察模型业绩在训练集和测试集上的变化。如果业绩变化很大，在参数取任何其他值时业绩都很糟糕，模型很有可能存在数据迁就偏差。

各种简化模型的方法都值得尝试。决定是否交易真的需要五个不同的

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

条件吗？逐个移除这些条件，模型在训练集上的业绩会在哪个临界点显著降低呢？更重要的是，移除这些条件时，模型在测试集上的业绩会相应降低吗？即使模型在训练集上的业绩可能会降低，只要没有显著降低，就应该尽可能地移除更多的条件、约束和参数。（但不能为了提升测试集上的业绩而增加条件和参数，或调整参数值，如果这样做了，就等于把测试集当成训练集使用了，并且有可能使模型再度产生数据迁就偏差。）

在简化了参数集和条件，并确保样本外测试的业绩在参数和条件微小变化时不受显著影响之后，应考虑将资金分配到不同的参数值和条件集。这种资金在参数上的平均化，将进一步确保模型的真实交易业绩与回测业绩不会相差太大。

3.5 交易成本

没有考虑交易成本的回测业绩是不真实的。第 2 章谈到了佣金、流动性成本、机会成本、市场冲击及滑价等交易成本，并且举了几个如何在策略回测中处理交易成本的例子。一项高夏普比率的策略，在考虑交易成本后变得无利可图，这是完全可能的，见例 3.7。

例 3.7　交易成本与均值回归模型

麻省理工学院的 Amir Khandani 和 Andrew Lo 提出了一个简单的均值回归模型（可参见 web.mit.edu/alo/www/Papers/august07.pdf）。策略非常简单：买入前一交易日日收益最差的股票，卖空前一交易日日收益最好的股票。在不考虑交易成本的情况下，1995 年以来，这一简单策略的业绩一直非常出色（2006 年的夏普比率为 4.47）。我们要研究的是，假定每笔交易成本为 5 个基点时，策略业绩如何变化。（此处的一笔交易是指一笔买入或一笔卖出，而不是一笔来回交易。）这个例子不仅向我们展

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

示了交易成本对业绩的影响，也展示了交易多只股票时 MATLAB 在模型回测中的强大功能。用 Excel 对模型在若干年中的交易进行回测需要设置一大堆标签，十分不便。但使用 MATLAB 也会遇到一个问题，即如何在几百个数据标签中检索历史数据，尤其是无存活偏差的数据。此处暂不考虑存活偏差的问题，因为获取无偏差数据十分昂贵，我们只要记得，所得的预测业绩相对实际业绩被高估。

对股票选择策略进行回测时，遇到的第一个问题永远是：哪些股票可用于回测？标准普尔 500 指数成分股通常是首选，因为它的流动性是最好的。标准普尔 500 指数的当前成分股可从标准普尔网站下载（www.standardandpoors.com）。由于成分股一直在变化，你下载的名单可能会与我的不同。你可在 epchan.com/book/SP500 20071121.xls 找到我的名单。下载所有这些股票的历史数据，最简便的方法是购买一份 HQuote Pro 的拷贝（可在 HQuote.com 下载）。这个软件可以使你很方便地剪切粘贴所需数据的所有标签列。使用软件检索和更新 2001 年 1 月 1 日至今的所有数据，只选择日期、开盘价、最高价、最低价、收盘价和成交量，不含数据名称栏，保存为文件“Export.txt”。通过以下 MATLAB 程序将其处理为运算时可使用的数据，在本地目录下保存为二进制文件“SPX_20071123.mat”。

```
clear;
inputFile='Export.txt';
outputFile='SPX_20071123';
[mysym, mytday, myop, myhi, mylo, mycl, myvol]=...
textread(inputFile, '%s %u %f %f %f %f %u','delimiter', ',');
% 找到不重复的股票集合
stocks=unique(mysym);
% 找到不重复的日期集合
tday=unique(mytday);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
op=NaN(length(tday), length(stocks));
hi=NaN(length(tday), length(stocks));
lo=NaN(length(tday), length(stocks));
cl=NaN(length(tday), length(stocks));
vol=NaN(length(tday), length(stocks));
for s=1:length(stocks)
stk=stocks{s};
% 找到当前股票的下标
idxA=strmatch(stk, mysym, 'exact');
% 找到当前日期集的下标
[foo, idxtA, idxtB]=intersect(mytday(idxA), tday);
% 获取当前股票的价格
op(idxtB, s)=myop(idxA(idxtA));
hi(idxtB, s)=myhi(idxA(idxtA));
lo(idxtB, s)=mylo(idxA(idxtA));
cl(idxtB, s)=mycl(idxA(idxtA));
vol(idxtB, s)=myvol(idxA(idxtA));
end
save(outputFile, 'tday', 'stocks', 'op', 'hi','lo', 'cl', 'vol');
```

接下来，用这些历史数据来回测不考虑交易成本的均值回归策略：

```
clear;
startDate=20060101;
endDate=20061231;
load('SPX 20071123', 'tday', 'stocks', 'cl');
% 日收益率
dailyret=(cl-lag1(cl))./lag1(cl);
% 等权重的市场指数收益率
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
marketDailyret=smartmean(dailyret, 2);
% 股票的权重与其离市场收益率的距离的负数成正比
weights=...
-(dailyret-repmat(marketDailyret,[1 size(dailyret,2)])). /
repmat(smartsum(isfinite(cl), 2), [1 size(dailyret, 2)]);
% 删除没有有效价格或日收益率的股票
weights(~isfinite(cl) | ~isfinite(lag1(cl)))=0;
dailypnl=smartsum(lag1(weights). * dailyret, 2);
% 删除考虑期外的盈亏
dailypnl(tday < startDate | tday > endDate) = [];
% 夏普比率应该是 0.25
sharpe=sqrt(252) * smartmean(dailypnl, 1)/smartstd(dailypnl, 1)
```

这个程序文件包括在 epchan. com/book/example3_7. m。注意，2006 年的夏普比率只有 0.25，而非原作者所报告的 4.47。业绩的显著降低，是因为本例的回测使用的是标准普尔 500 指数所有成分股。如果读了作者的文章，你就会发现绝大部分收益来自小盘股和微盘股。

在上面的 MATLAB 程序中，我使用了三个新函数“smartsum”、“smartmean”和“smartstd”，其功能与“sum”、“mean”和“std”三个常用函数的功能相似，但能自动忽略数据中所有 NaN 元素。这些函数在进行回测时非常有用，因为股票的价格序列经常是断断续续的。这些文件在 epchan. com/book 都能找到。

"Smartsum. m"

```
function y = smartsum(x, dim)
% y = smartsum(x, dim)
% 对 dim 维求和,忽略 NaN
hasData=isfinite(x);
x(~hasData)=0;
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
y=sum(x,dim);
y(all(~hasData, dim))=NaN;
```

"smartmean. m"

```
function y = smartmean(x, dim)
% y = smartmean(x, dim)
% 对 dim 维求均值，忽略 NaN
hasData=isfinite(x);
x(~hasData)=0;
y=sum(x,dim)./sum(hasData, dim);
y(all(~hasData, dim))=NaN; % 若 y 的所有元素都是 NaN，则 y 取值为 NaN
```

"smartstd. m"

```
function y = smartstd(x, dim)
%y = smartstd(x, dim)
% 对 dim 维求标准差，忽略 NaN 和 Inf
hasData=isfinite(x);
x(~hasData)=0;
y=std(x);
y(all(~hasData, dim))=NaN;
```

现在考虑每笔交易成本为 5 个基点时的回测。

```
% 减去交易成本后的盈亏
onewaytcost=0.0005; % 假设交易成本为 5 个基点
% 删除考虑期外的权重
weights(tday < startDate | tday > endDate,:) = [];
%权重变化时产生交易成本
dailypnlminustcost=...
dailypnl - smartsum(abs(weights-lag1(weights)), 2).* onewaytcost;
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 夏普比率应该为 -3.19
sharpeminustcost=sqrt(252) * smartmean(dailypnlminustcost,1)/...
smartstd(dailypnlminustcost, 1)
```

现在，这一策略的业绩非常差！

3.6 策略改进

有一些常用方法可用来改进初次回测业绩不佳的策略。如何在改进策略时不引进数据迁就偏差，并保持少参数的简单模型，更像是艺术而非科学。一个与参数优化相同的指导原则是：任何策略改进要同时提高训练集和测试集的业绩。

经常有一些非常简单的策略，为交易员所共知，收益在不断下降，但仍然盈利。股票配对交易就是一个例子。收益下降的原因在于众多交易员会利用这种套利机会，从而逐渐消除盈利。然而，通常会对基本策略进行微小调整，来提升收益。

这些微小调整往往不如基础策略那样为人熟知，因此交易员也很少采用。比如有时会排除某只或某组特定股票。例如，交易员会倾向于从技术交易程序中排除价格易受消息影响的医药股，或面临并购的股票。还可以改变进出市场的时间或交易频率，甚至简单到选择一组不同的股票。如例 3.7 就可以看到，一项策略应用于小盘股时夏普比率很出色，而应用于大盘股时却非常糟糕。

策略的改进，最好基于经济学基本原理，或者透彻研究过的市场现象，而不是依据一些主观的试错法则。否则，就有可能产生数据迁就偏差。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例 3.8 清仓策略的微小调整

我们改进一下例 3.7 中的均值回归策略。考虑交易成本后，2006 年策略的夏普比率从 0.25 变为 −3.19。对策略仅做一下改动：在市场开盘而非收盘时更新头寸。将 MATLAB 代码中所有的“cl”改为“cp”。

你会发现，不考虑交易成本，夏普比率会增加到 4.43；考虑交易成本后，夏普比率也增加到有利可图的 0.78！用标准普尔 400 中盘股指数和标准普尔 600 小盘股指数的成分股进行回测，会进一步提高夏普比率，这留给读者自行练习。

3.7 小 结

策略业绩的真实历史模拟过程叫做回测。回测的依据在于假设策略的未来业绩会与过去业绩一致，不过你的投资经理会不厌其烦地告诉你，没有人能保证一定会这样。

在进行真实历史回测，以及减少未来业绩偏离回测业绩时，会牵涉许多细节问题，包括：

- 数据：分拆和股息的调整，日最高、最低价的噪声，存活偏差。
- 业绩度量：年化夏普比率和最大挫跌。
- 前视偏差：在过去的交易决策中使用无法得到的事后信息。
- 数据迁就偏差：拟合历史数据时使用过多参数，可以用大样本数据、样本外测试和敏感性分析来避免此类偏差。
- 交易成本：交易成本会影响策略业绩。
- 策略改进：通过策略的微小调整来优化业绩的常见方法。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

通过本章的学习，并做了例子和练习之后，你就有了一些实践经验，知道了如何检索历史数据以及使用EXCEL或MATLAB回测策略。

测试策略时，由于时间和其他条件的限制，不可能避免所有陷阱。这种情况下，可以略过一些预防措施，快速感受一下这个策略是否具有潜力，是否值得做进一步检验。有时，最严格和仔细的回测都无法查验出来的问题，通过几个月的仿真交易或真实交易就会显现出来。在模型应用过程中，会经常遇到这些问题。

一旦对策略进行回测并得到合理业绩后，就为下一步开展交易事业做好了准备。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第4章 创建交易业务

本章暂时不涉及量化交易的技术层面，而将注意力放在业务层面。如果你的目标是保持独立交易员的身份，不为任何资产管理机构工作，交易的业务结构的选择便十分重要。首先，选择开立零售经纪账户，还是加入自营交易公司。其次，确定零售经纪商或自营交易公司的特征，对你而言非常重要。最后，为了执行量化交易策略，还要决定需要哪些设备。

4.1 业务结构：零售还是自营？

作为一名量化交易员，可以选择是完全独立交易还是半独立交易。若是选择完全独立交易，那么只需开立一个零售经纪账户，然后存入一定现金，便可开始交易。没有人会质疑你的策略，也没有人会对你的交易指手画脚。然而，如果持有隔夜头寸，SEC 的规则 T 会限制你的杠杆，大约是净值的两倍。当然，你要自负盈亏。

此外，你也可以选择加入诸如 Bright Trading、ECHOtrade 或 Genesis Securities 等“自营交易公司”，成为其会员，不过必须通过美国

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

国家证券交易商协会（NASD）的系列 7 考试，取得注册证券经纪代表资格。你仍然需要用自己的资本开立一个账户，但与零售账户相比，可以获得更高的杠杆（或“购买力”）。你能分得的收益比例，取决于你投入的自有资本比例。尽管有负债，你的亏损仅限于初始投资。（事实上，如果设立一家小型公司或有限责任公司，并通过这个实体开立零售经纪账户，也是有限责任。）通常，你会受到自营交易公司的培训，这可能会产生额外成本。除了必须遵守 SEC 和 NASD 制定的法规外，还要遵守自营交易公司对会员的规定和规章。

自营交易公司的规定和规章，听起来不是什么好事。但事实上，有些规定（如禁止交易低价股票或禁止持有隔夜空头头寸）其实是一种自我保护的风险管理措施。在市场走强时，交易员经常会抱怨这些规定限制了他们的交易灵活性和盈利水平。他们甚至可能会开立零售交易账户进行独立量化交易。然而，在遭受重大挫跌时（几乎是不可避免的），又希望有人能限制一下他们的风险偏好，并且开始后悔这种不受约束的自由。

开设零售账户还是加入自营交易公司，通常取决于你的资本需求、策略风格和交易水平。例如，对于低风险的市场中性策略，需要一个远超规则 T 允许范围的高杠杆以产生好的收益，就该选择自营交易公司。然而，对于不需要太多资本的高频期货交易，零售经纪商能节省很多成本并减少不必要的麻烦。类似的，善于管理风险、控制情绪的老练交易员，可能并不需要自营交易公司提供指导，但缺乏经验的交易员却能从中受益。

如果你发现了独特的、高盈利策略，就会有另一层的考虑。在这种情况下，应该开立零售交易账户，因为如果通过自营交易公司的账户进行交易，自营交易公司便会知晓你的高盈利策略，也许会用大量自有资本“搭便车”。此时，随着时间的推移，你的策略将承受更多的市场冲击交易成本。

表 4—1 概括了零售经纪商和自营交易公司的优缺点。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

表 4－1　　**零售经纪商和自营交易公司的比较**

事项	零售经纪商	自营交易公司
开户要求	无	需通过 NASD 系列 7 考试并受其他 NASD 规则约束
初始投资	较多	较少
可用杠杆	SEC 规则 T。一般两倍于隔夜头寸或四倍于当日头寸	公司自行决定。日头寸和对冲头寸可高达 20 倍甚至更高
亏损责任	无限责任。除非此账户由小型公司或有限责任公司开立	仅限于初始投资
佣金和费用	较低的佣金（每股不到 0.5 分钱）。每月费用少	较高的佣金。每月费用较多
经纪商破产风险	无风险。由证券投资者保护公司（SIPC）作保	有风险。账户无保险
培训、监控、指导	无	能提供。但可能要收费
交易机密泄露	很少或无风险。特别是在零售经纪商没有自营交易业务时	有风险。管理人能够轻松搭上盈利策略的便车
交易风格限制	无，只要不违反 SEC 的规定即可	可能有，如禁止持有隔夜空头头寸
风险管理	自己负责	公司负责

最后提醒一点：有些人可能认为加入自营交易公司会有税收优势，因为任何交易亏损都可以从现金收入中扣除而不产生资本亏损。实际上，即使你开立的是零售经纪账户，也可以申请交易员税种，这样交易亏损也可以抵扣其他收入，而不仅仅是其他资本收益。关于交易业务的税收考虑，可参见 www.greencompany.com。

4.2　选择一家零售经纪公司（或自营交易公司）

许多交易员在选择经纪商或自营交易公司时，只考虑一条标准：佣金

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

比率。这条标准的确很重要，因为假若策略的收益率很低，高佣金可能会使得这种策略无利可图。但是，还有许多其他因素需要考虑。

其实佣金只占总交易成本的一部分，有时甚至比例会很小。经纪商的成交速度和是否提供所谓的“暗池”流动性，也要算入交易成本。交易所之外的机构指令，或经纪商客户之间的内部交叉订单，构成暗池流动性。这些指令并不显示在买卖报价之中。有些“另类交易系统”（如 Liquidnet 和 ITG's Posit）会提供暗池流动性交易。你的经纪商可能会使用一家或多家外部提供商的系统，也可能只使用内部交叉网络，甚至不使用“另类交易系统”。

有时，有实力经纪商的先进交易系统及其提供的高速暗池流动性，使指令以最优价格成交，从而足以弥补其高额佣金。不过，只有同时通过多个经纪商进行交易，并比较真实的交易成本时，才可能对成本和收益进行权衡。

例如，我一般使用高盛的 REDIPlus 交易平台进行交易，这个交易平台的 Sigma X 执行引擎会同时对内部交叉网络和外部流动性提供商传递指令。我发现，此平台为我撮合的每股交易价格通常会比 Interactive Brokers 提供的价格高那么几美分，这足以支付高盛相对较高的佣金了。

另一个需要考虑的因素是可交易品种的范围。许多零售经纪商或自营交易公司不允许进行期货或外汇交易。这将严重制约交易业务的增长。

对于量化交易员而言，紧接以上两个普遍性标准之后的一个重要标准是：交易平台是否提供应用程序接口（API）。通过 API，交易软件能够接收实时数据，产生指令，并自动发出指令以执行。第 5 章将详细讨论 API。此处唯一需要知晓的一点是，如果没有 API，就不可能进行高频量化交易。

与 API 密切相关的是仿真交易账户。如果经纪商不提供仿真交易账户，很难在不冒真实亏损风险的前提下测试 API。我所知道的可以提供仿真交易账户的经纪商有 Interactive Brokers、Genesis Securities、PFG

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

Futures（期货交易）和 Oanda（外汇交易）。

除仿真交易账户外，有些经纪商也会提供“模拟机”账户（如 Interactive Brokers 的演示账户）。在模拟机账户上，历史报价被当作实时报价来显示，同时自动交易程序能够根据这种报价随时进行调试。

最后，自营交易公司的声誉和财务状况也很重要。若选择零售经纪商则不必考虑这些，正如表 4－1 所示，零售账户由证券投资者保护公司(SIPC) 作保，而自营账户则没有。因此，自营交易公司漂亮的资产负债表和过硬的风险管理措施就很重要，可以有效防止其他会员交易员的亏损而导致的破产（世通公司和瑞富公司的破产就是很好的例子）。同时，还要确认这家公司是在交易所注册的经纪自营商，这样它就得接受 SEC 和交易所的定期审计（就在本书撰写之时，非注册经纪自营商的自营交易公司可能会被 SEC 全部关闭，Tuco Trading 已于 2008 年 3 月第一个被关闭）。当然，对于一个门外汉来说，评估这些自营交易公司是否具备上述优秀特征是比较困难的，但你可以根据在线专栏 www. elitetrader. com 上，曾在或正在这类公司工作的交易员所发表的观点来了解它们的声誉。

如果对该开立零售经纪账户还是自营账户心存犹疑，或对选择哪家零售经纪商或自营交易公司举棋不定，可以两个账户都开立，甚至开立多个账户。与自营交易公司的全职雇员不同，只作为会员，特别是远程访问会员，通常不会强制你签署非竞争协议。你可以成为不止一家自营交易公司的会员，也可以同时拥有零售和自营交易账户，只要将其作为“外部交易活动”对所有这些自营交易公司和 NASD 充分披露并取得事先许可即可。当手握多个账户时，就可以很轻松地了解哪种成本结构更加有利可图，哪个账户可以为自动交易系统提供更好的交易设施和交易工具。实际上，每个账户都各有优缺点，不同的策略可以用适合其特点的不同账户来交易。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

4.3 设　备

现在，你已经为你的交易业务建立了法律和管理结构，该是考虑设备的时候了。许多自营交易公司允许其会员在家中进行远程交易，不论是零售交易员还是自营交易员。如果你是一名不需要账户管理指导的自营交易员，且对建立设备的能力很有信心，那就没理由不进行远程交易。

在业务的起步阶段，设备可以比较简单。你的家庭办公室可能已经都有了：一台性能良好的个人电脑（任何双核的新电脑即可），高速网络，以及防中断电源（UPS），后者可以使你的电脑不会因为电力故障而在交易过程中突然关机。购买所有这些设备最多不超过 1 000 美元，如果没有订购有线电视，每月的花费不会超过 50 美元。

有些交易员可能会订购 CNBC 或 CNN 频道。尽管这没什么坏处，但许多专业量化交易员会觉得没必要，他们会订购另一类专业实时新闻工具，如汤森路透、道琼斯或彭博。相比于彭博每月将近 2 000 美元的高昂费用，汤森路透和道琼斯每月只需花费 100～200 美元（尽管可能需要签订包年合同）。彭博还有一个免费的互联网广播（www.bloomberg.com/tvradio/radio）用来播送最新的重大新闻和事件评论。也可以订购 CBNC PLUS 来随时向你的电脑推送新闻视频，而不必在办公室安装一台电视。当然，接收过多的即时信息可能并不会使你赚到更多钱。例如，美盛集团的迈克尔·莫布森做了一项研究，发现那些在赛马比赛时接收过多信息的新闻记者，其预测的成功率都很低（参见 *Economist*，2007a 或 Oldfield，2007）。

随着交易增多，可以逐步进行设备升级。如购买一台速度更快的电脑，目前四核电脑已经算是高配了。当然，在你读到本书时，也许八核或更好的个人电脑也已变得十分普及。根据《纽约时报》的一篇文章

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

(Markoff，2007)，八核个人电脑在 2010 年上市。

你可以把多个屏幕与同一台电脑连接起来，这样就相当于扩大了屏幕的可视空间，便于监控不同的交易程序和多个投资组合。还可以考虑把网络升级为 T1 线。第 2 章说过，指令传送给经纪商过程中的任何延迟都会导致滑价，从而造成亏损。在毫秒级的高速变化市场中，网速升级是很费钱的。(T1 线的成本在 700～1 500 美元之间，而 cable 和 DSL 通常不到 50 美元。T1 线的信息传输速度是每秒 1.5 兆比特，大概是 cable 和 DSL 的两倍。)

一旦完成策略的最终测试，且发现在实际交易中表现不错，就该是扩大交易规模的时候了。还得同时考虑所谓的“业务持续计划”，使交易不受网络中断、电源中断、洪水等自己家中可能发生的灾难的影响。可以把交易程序安装在托管商的远程服务器上，甚至可以把安装了交易程序的电脑放在托管商那里。(这种程序托管或机器托管每月将花费数百到上千美元不等。) 你可以通过普通的远程应用程序（如每月 15 美元的 GotoMyPC）来监控、维持或升级这些在远程服务器上运行的交易程序，远程服务器直接将指令发送给经纪商。这样做的优点在于，交易几乎不会被迫中止，而且托管商的网速要比家里或办公室更快。实际上，对于超高频交易而言，把服务器安放在离指令成交的交易所越近越好。在 Google 搜索关键词“服务器托管”或“服务器配置”，就能找到许多提供这种服务的托管商。

4.4 小 结

本章关注连接量化交易的研究阶段和执行阶段之间的相关决定和步骤，讨论了零售交易和自营交易各自的优缺点，以及选择一家零售经纪商或自营交易公司时应该考虑的问题。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

简而言之，零售经纪商可以给予完全的自由和更好的资本保护，但杠杆较低；自营交易公司给予的自由和资本保护较少，但杠杆较高。找到一家合适的零售经纪商相对容易，我就只用了不到一个月时间的调查并在一家零售经纪商开立了账户，至今也没想过要换一家。找到一家合适的自营交易公司则要麻烦一些，因为要签合约并通过系列 7 考试。我大概用了几个月的时间才选择了一家并开立了账户。

当然，你可以同时拥有零售和自营交易账户，以满足不同策略的特殊需求。这样也可以很容易地比较成交速度和流动性的深度。

无论是选择零售经纪商还是选择自营交易公司，都要确保交易账户和系统满足以下特征：

- 相对较低的佣金
- 可交易金融工具品种广泛
- 有足够深度的流动资金池
- 最重要的是，获取实时数据和传送指令的 API

本章还谈到为满足交易业务需要逐步升级的设备。交易员的操作环境包括：

- 一台双核或四核电脑
- 高速网络（cable、DSL 或 T1）
- 防中断电源
- 实时数据和新闻来源
- 服务器托管

事实上，配置交易设备是相当简单的事情，因为你的家庭办公室可能之前已经全部都有了。交易上百万美元投资组合所需的设备，仅需几千美元的初始投入，以及每个月几百美元的运营成本。但如果你打算提升交易能力或提高收益，就需要额外的购置设备投资了。

在完成这些步骤后，就要构建能执行策略的自动交易环境了，下一章将讨论这个问题。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第5章 交易执行系统

到此，你已完成了策略回测（如例 3.6 中的配对交易策略），选定了经纪商（如 Interactive Brokers），配置了操作环境（刚开始只需要一台电脑、高速网络和实时新闻来源）。当生成和传送指令给经纪商的自动交易系统（ATS）设置完成后，执行交易策略的准备工作便基本就绪。本章讨论如何建立这样的自动交易系统，并讨论交易成本以及预测业绩与回测的偏差最小化的方法。

5.1 自动交易系统的功能

自动交易系统从经纪商或其他数据提供商那里获取最新的市场数据，运行交易算法生成指令，传送指令给经纪商以执行。有时，所有这些步骤都是自动的，就像电脑里安装的一个桌面应用程序一样。有时，只有部分过程是自动的，有些步骤需要手工操作。

全自动交易系统的优势是可将人为错误和延迟降到最低。对于高频交易系统，全自动交易系统是必不可少的，因为任何人工操作都会产生足以

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

严重影响业绩的延迟。但是，全自动交易系统十分复杂且昂贵，通常需要专业程序员掌握 Java、C# 或 C++等高性能程序语言，以实现与经纪商的应用程序接口（API）对接。

对于低频量化交易策略，半自动交易系统即可：用 Excel 或 MATLAB 等程序生成指令，然后用经纪商提供的现成工具（如组合交易器或差价交易器）来传送指令。如果经纪商提供连接 Excel 的动态数据交换功能（参见下文），就可以在 Excel 中编写一个宏，通过运行宏将指令传送给经纪商。这样，就不需要用复杂的程序语言来编写应用程序。但是，这也就意味着你不得不完成一些手工操作，以备传送指令。

无论是半自动交易系统还是全自动交易系统，通常都需要输入一些经纪商或其他数据提供商不提供的数据。例如，实时数据流通常不提供的盈利预测或分红数据。很多网站都会免费提供这些非价格数据，不过通常是 HTML 格式，不能直接使用。因此，自动交易系统必须能够在这些网页上检索、读取数据，并将其转换为策略可以使用的格式。这种网页检索和读取程序，可以轻易地安装到 MATLAB（见第 3 章例 3.1），以及 Perl 之类的脚本语言里。

接下来，我们将讨论这两种交易系统的细节。我也会介绍如何聘请编程顾问，因为你也有可能找人帮忙来实现交易策略的自动化。

5.1.1 建立半自动交易系统

半自动交易系统中（如图 5－1 所示）一般用 Excel 或 MATLAB 等好用的软件生成指令。生成指令的程序通常就是回测程序，毕竟，所执行的量化交易策略就是回测的策略。当然，要记得更新输入的数据文件，以反映最新的数据。这项工作要么由能够直接进入网站检索所需数据的 MATLAB 完成，要么由一个独立的程序来完成，如前面提到的 HQuote。在下面的例子中，MATLAB 仅用作将这些数据转换为合适的格式，以便策略程序生成指令。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

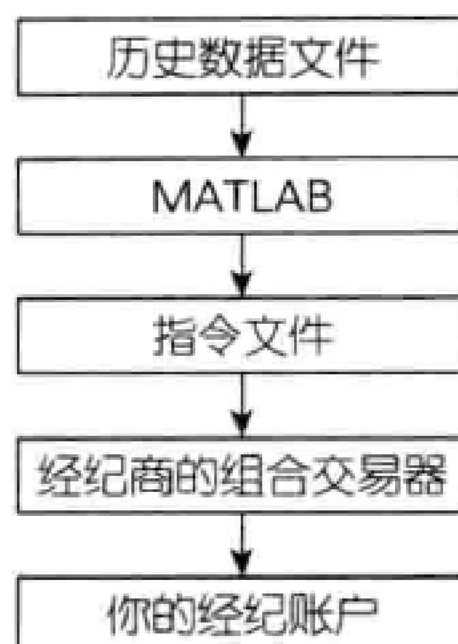


图 5—1 半自动交易系统

若最新数据是前一天的收盘价，数据更新比较容易，若是当天最新成交价，就比较困难。如果所需数据是最新成交价，经纪商或数据提供商就必须提供一个能够自动更新 Excel 输入文件的动态数据交换（DDE）链接。绝大多数为交易员提供服务的经纪商，都会提供这种动态数据交换链接，如 Interactive Brokers，Genesis Securities 和高盛 REDIPlus 等。许多自营交易公司会使用其中一家经纪商来执行交易，因此，你可以获得这些经纪商提供的全部实时数据和指令输入技术。

动态数据交换链接可以直接导入 Excel 电子表格，自动将实时数据写入单元格内。不同经纪商的动态数据交换链接的格式不尽相同，但看起来都跟这个差不多：

=accountid | LAST! IBM

其中，LAST 表示最新价格，IBM 是股票代码。

可通过运行 Excel 宏（与表格关联的 VB 程序）或 MATLAB 程序来生成指令，它们可以检查表格中的信息和价格，运行交易算法，将指令写入另一个文本文件，文件每一行包含三项内容：股票代码、方向、数量。例如：

(" IBM", " BUY", " 100")

就可能是输出指令文件中的一行。有时，经纪商还需要其他信息，如当天有效或撤单前均有效。所有这些补充信息要写入指令文件的每一行。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

在生成包含指令的文本文件后，就可以将指令文件上传给经纪商的组合交易器或差价交易器，以备传送。

组合交易器是可以向经纪商一次上传多只股票多个指令的应用程序。差价交易器是可以指定多对股票或证券及各对指令触发条件的应用程序。差价交易器在整个交易日中，能监控实时价格并检查是否满足触发条件。如果经纪商的动态数据交换链接允许传送指令，也可以运行一个 Excel 宏文件来扫描指令文件，并且可以将所有指令一次性传送到账户。

我的两个经纪商 Interactive Brokers 和 REDIPlus，都有组合交易器和差价交易器，以及可以更新数据和传送指令的动态数据交换链接。(Interactive Brokers 的差价交易器只能用于期货日历差价。股票差价可通过“通用组合”在某时刻指定股票配对来实现，不过期货和股票都有指令触发条件的持续监控。)

这里，介绍一下我是如何使用 Interactive Brokers 的组合交易器。每天开盘前，我运行一个 MATLAB 程序检索市场数据，运行交易算法，并将一组指令写入指令文件，指令文件往往超过 1 000 行（相应地，超过 1000 个标签)。之后，我在交易页面启动组合交易器，并用组合交易器一次向账户上传包括所有指令的指令文件。有些指令可能在开盘时就被执行了，有些指令可能会在稍后被执行，有些可能没被执行。收盘前，我要点击按钮取消所有未执行的指令。最后，如果我想将所有头寸清仓，只需点击组合交易器里的另一个按钮生成相应的清仓指令即可。

我用 REDIPlus 的差价交易器来执行例 3.6 中的配对交易策略，因为可以在交易日的任何时候通过差价交易器下达指令，而无需等到收盘。同样，我在开盘前用 MATLAB 收集市场数据，运行配对交易算法，并写入所有配对股票的限价。(注意，这里的限价是对差价的限定，而不是单只股票。如果是限定单只股票，普通限价指令就能做到，不需要差价交易器。) 之后，在已包含之前所选定配对的差价交易器中，我根据 MATLAB 的输出结果人工调整限价，其中交易器已经包含了我之前指定

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的所有配对股票。（实际上，这个步骤也可以自动完成：MATLAB 可以把所有差价指令信息写入一个 Excel 文件，并上传到差价交易器。）点击另一个按钮就可以开始全天候监控股票价格和指令执行信息了。

我也会用 REDIPlus 的动态数据交换链接来传送另一个组合交易策略的指令。我先用 MATLAB 在每个 Excel 单元格里生成一个恰当的动态数据交换链接公式，这样每行标签下的数据就可以自动更新了。开盘后，我通过电子表格的宏检查每个数据标签并将其（连同电子表格中的其他指令信息）上传到 REDIPlus 的账户中。

一般来说，半自动交易系统适用于每天只需运行少数几次来生成一波或几波指令。即使经纪商的应用程序接口（API）提供用 Excel VB 宏传送指令的功能，要频繁运行程序来获取最新数据并生成一波又一波的指令，程序的速度会非常慢。这种情况，就必须建立全自动交易系统。

5.1.2　建立全自动交易系统

全自动交易系统（见图 5—2）可在整个交易日，不间断地反复运行交易算法，持续监控最新价格并不断生成指令。指令通过应用程序接口（API）自动传送到经纪账户，因此，无需在组合交易器或差价交易器中完成交易，也无需手动运行 Excel 电子表格的宏。要做的全部工作就是在每天早上点击“开始”按钮，并在每天晚上点击“关闭”按钮，程序会自动完成全部交易。

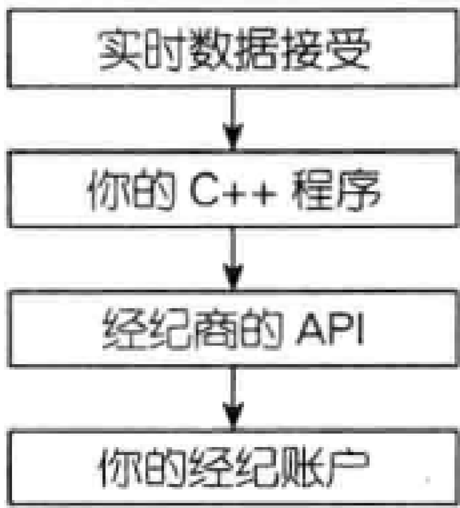


图 5—2　全自动交易系统

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

全自动交易系统的运行需要经纪商提供 API，用于检索数据和传送指令。经纪商提供的 API 通常适用于 VB、Java、C# 或 C++这些常见的编程语言，因此，自动交易系统也必须是用这几种语言编写的。很遗憾，据我所知，没有经纪商提供适用于 MATLAB 的 API。因此，不能用 MATLAB 构建自动交易系统中的指令传送模块。

从理论上讲，可以用 Excel 电子表格和内置的宏来构建全自动交易系统，所要做的就是在宏中建立一个循环，这样就可以每天在适当的时间用动态数据交换链接更新单元格数据并传送指令。遗憾的是，通过动态数据交换链接进行数据升级的速度很慢，并且经纪商通常会限制每次更新的股票数量。(除非你在上一个交易月贡献了大笔佣金，Interactive Brokers 只允许更新 100 只股票的数据。) 同样，通过动态数据交换链接传送指令的速度也很慢。考虑到策略要反映一天中实时市场数据的变化，用电子表格来构建全自动交易系统并不可行。

TradeStation 等经纪商可以提供进行完整回测和指令传送的平台。如果你使用这样的平台进行回测，只需对平台进行一定的配置，程序就能够将真正的指令发送到你的账户上了。这样，无论是进行回测还是自动执行交易，都不需要编写自己的软件了。但就像第 3 章提到的，这种自营系统的缺点是，在构建策略时不如 MATLAB 这样的多用途程序语言那么灵活。比如，如果要跟踪一个基于主成分分析的复杂的数学策略（如例 7.4），使用 TradeStation 来做回测是非常困难的。Alphacet' s Discovery 等更加先进的集合交易平台的确可以为回测和执行提供更多的算法，但价格也不是普通的独立交易员所能承受的。

聘请编程顾问

建立自动交易系统比进行策略回测需要更多的专业编程技巧。对于高频交易策略，交易执行速度是非常关键的。寻找一个编程顾问比自己建立交易执行系统要轻松不少。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

聘请编程顾问并不需要花很多钱。请一个有经验的程序员，每小时一般是50～100美元。有时可以事先就整个项目谈定一个固定价格，以我的经验，绝大多数独立交易员的编程项目在1 000～5 000美元。如果开户的经纪商提供API，经纪商通常会推荐一些熟悉它的API的程序员。（例如，Interactive Brokers就有一个网页，编程顾问可以在上面提供服务信息。）也可以在elitetrader.com上寻找编程顾问（或发布需求信息）。最后一招，可以在craigslist.org上看到自由程序员发布的成百上千条自荐广告。但我觉得craigslist.org上自由程序员的水平参差不齐，尤其是他们对金融市场和交易技术的了解不多，而这对于成功构建自动交易系统是很关键的。

聘请程序员还需要考虑一个问题：策略如何保密？当然，你可以和他们签订保密协议（可在许多法律文件网站上免费下载），但在交易系统构建完成后，就算他们真的窃取了你的策略并在其个人账户中运行，也是几乎无法发现的。有几种方法可以解决这个问题。

首先，我之前提到过，绝大多数你认为是独创的策略其实很多有经验的交易员都知道。所以，不论你是否愿意，其他人早已进行过相似策略的交易，并影响了你的收益。再有一两个交易员参与进来也不会对收益产生更多的影响，除非他是机构基金经理。

其次，如果你的策略交易量巨大（如绝大多数期货交易策略），你聘请的那位没有道德的编程顾问根本不会对你产生什么影响。

最后，你可以把编程工作拆分给不同的人去完成，也就是说，可以聘请不同的程序员分别完成自动交易系统不同部分的构建。如让一名程序员编写可用于不同策略的执行程序，让另一名程序员编写尚未输入参数的策略。这样，第一名程序员不知道策略，第二名程序员没有执行策略所需的基础程序，并且两名程序员都不知道你的策略所使用的真实参数值。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

5.2 最小化交易成本

第 3 章显示了交易成本如何影响策略的实际收益。除了选择佣金较低的经纪商或自营交易公司外，还可在交易执行方法上做些文章来降低交易成本。

为降低佣金，你可以避免交易低价股票。一般来说，机构交易员不会交易任何一只价格低于 5 美元的股票。低价股票不仅会增加总佣金成本(对于一定的资本，需要买进或卖出更多股票)，还会有相对较高的买卖差价，从而增加总流动性成本。

为减小市场冲击成本，应根据股票的流动性来限制指令规模（股票数量)。衡量流动性的一个常用指标是平均日交易量（要平均的回溯期长短由你自己决定)。根据经验规则，单个指令的股数不宜超过平均日交易量的 1%。作为一名独立交易员，你可能认为达到 1%的门槛并不容易，对于标准普尔 500 指数的大盘股，你的想法可能是对的。但是，你也许会对一些小盘股的低流动性感到惊讶。

例如，在我写本书的时候，IRN 是标准普尔 600 小盘股指数的成分股。其 3 个月的平均日交易量约是 51 000 股，今天的收盘价是 4.45 美元。因此，平均日交易量的 1%只有 510 股，仅值 2 269 美元！

另一种降低市场冲击的方法是，根据股票的市值来决定指令规模。决定指令规模的方法并不是精确的科学，但绝大多数从业人员都不建议采用线性比例方法，因为不同上市公司的市值差别很大，从几千万到几十亿不等。线性比例方法（例如按市值的一定线性比例来确定投入的资本）会导致交易组合剔除掉绝大多数小盘股和微盘股，这就失去了分散投资的好处。如果我们用线性比例方法，最大大盘股的资本权重会是最小小盘股的 10 000 倍。为了获得分散化投资的好处，在满足上面所说的流动性（交

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

易量）条件下，这个权重比例不宜超过 10 倍。也就是说股票的资本权重与其市值的四次方根成正比。

还有一种减少市场冲击的方法。机构交易员会将大额指令拆分为许多小额指令并分开执行。这种方法确实能够减少市场冲击，但会导致另一种交易成本，即滑价。正如第 2 章所讨论的，滑价是触发交易信号的价格与全部指令平均执行价格的差。由于指令是在一段时间内被执行，滑价可能非常大。因为这种减少市场冲击的方法会增大滑价，不太适合指令规模通常不是很大的零售交易员。

但是，有时候滑价不是你能够控制的，也许是由于经纪商的执行速度太慢，这可能是软件原因（处理指令的软件太慢），风险控制原因（指令传递至交易所前，检查你账户是否有购买能力、各项风险控制是否达标），或网速原因（经纪商连接交易所的速度），或者也有可能是你的经纪商没有足够的“暗池”流动性。正如第 4 章所说，这些执行成本和问题会影响对经纪商的选择。

5.3 用仿真交易测试交易系统

建成自动交易系统后，用经纪商提供的仿真交易账户来测试一下是个不错的主意。仿真交易有许多好处，最主要的好处是，这事实上是在不亏钱的情况下能够查找自动交易系统软件漏洞的唯一方法。

通常，在一开始进行仿真交易时，就能发现策略中的前视偏差——在输入一条指令之前，可能没有办法去获得本应该获得的一些关键数据片段！如果发现了前视偏差，只能从头再来了。

你该运行自动交易系统，执行仿真交易，然后对仿真交易的盈亏与用最新数据进行的回测的理论盈亏进行比较。如果两者之差不是由交易成本引起的（包括执行仿真交易时的预期延迟），很有可能就是软件有漏洞。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

(第 4 章给出了一些提供仿真交易账户的经纪商名录。)

仿真交易的另一个好处是，可以使你对策略有更好的直观理解，包括盈亏波动、所用资本量、每天交易次数和包括数据在内的各种操作难题。你能够在回测中对策略的这些特征进行理论上的了解，但只有在每天持续不断的交易中才能够获得直观感受。回测也反映不出操作上的问题，如在每日开盘前下载全部所需数据的速度，以及如何优化真实交易中的操作流程。

(不要低估开盘前准备指令所需的时间。我每天早上大概要花 20 分钟的时间下载、分析所有历史数据，再花 15 分钟左右时间将所有指令传送到我的账户。如果你的交易策略需要开盘前 35 分钟之内的数据或新闻，你要么另建一个交易环境，要么修改策略。不进行仿真交易，很难测算完成这些事项所需时间。)

由于仿真交易是纯粹的样本外测试，如果仿真交易系统可以运行一个月，将极有可能发现数据迁就偏差。但有这种耐心的交易员不多，因为需要关注的事情还有很多（如正在运行的真实交易程序)。由于操作中的错误和疏忽，所以不关注，会导致仿真交易系统的业绩很差。因此，只有使用少量资金进行真实交易时，才能发现数据迁就偏差。

5.4 实际业绩偏离预期的原因

终于，经过辛勤测试和准备之后，你输入了第一条指令并被执行！不管是赚钱还是亏损，你明白，要判断交易业绩是否与预期吻合，仍为时过早。但如果一个月、两个月甚至一个季度后，策略业绩依然不佳甚至亏损，怎么办？对量化交易的新手来说，这种令人沮丧的经历再正常不过了。这时候就该找出是什么原因导致了实际业绩与预期的偏差。先从最简单的诊断开始：

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

- 自动交易系统软件是否有漏洞?
- 自动交易系统生成的交易与回测程序生成的交易是否匹配?
- 执行成本是否远高于预期?
- 是否交易了流动性差的股票从而导致了市场冲击?

如果执行成本显著高于预期，就需要好好回顾一下“最小化交易成本”一节了。

在完成上面的简单诊断后，就要面对造成偏差的两个最令人头痛的原因：数据迁就偏差和状态转换。

为了观察数据迁就偏差是否造成真实交易的糟糕业绩，可以尝试尽可能地减少策略中的规则和参数。如果减少后回测业绩完全崩溃，这种偏差就是存在的，就要放弃这个策略。如果回测业绩仍然合理，交易业绩不佳很有可能是运气不好。

状态转换是指金融市场结构或宏观经济环境发生了巨变，从而使得原本盈利的策略现在不盈利了。

近年来，两次明显的与市场（或监管）结构相关的状态转换，对某些策略造成了影响。

一次是股票价格采用十进制报价。2001 年之前，美国股票价格采用 1/16 美分或 1/18 美分的整数倍进行报价。2001 年 4 月 9 日开始，所有美国股票都采用十进制报价。这个看似无关紧要的变化，对市场结构产生了剧烈的影响，特别是对统计套利策略的盈利能力产生了负面影响。

要解释其中的原因，可能需要专门写一本书。简单来说，十进制报价减少了价格发现过程中的摩擦，而统计套利者与做市商一样，是从这个过程中的摩擦和低效来获利的。(这是 2008 年 Andrew Sterge 博士在哥伦比亚大学“所有的统计套利盈利都到哪儿去了?”的金融研讨会上给出的解释。其他机构从业人员在私下里也对我讲了相同的观点。）因此，我们可以预期，统计套利策略 2001 年之前的回测业绩要远远好于现在的业绩。

另一次状态转换关系到卖空股票。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

2007 年之前，SEC 规定，除非股票处于“涨点”或“零涨点”，否则不允许卖空。因此，如果回测用了在此之前的数据，很有可能由于缺少涨点而不会建立一个非常有利可图的空头头寸，或者以较大滑价建仓。2007 年 6 月，SEC 取消了涨点规则。因此，卖空策略的回测结果相比于 2007 年之前的真实交易业绩有一个人为的抬高。

事实上，除了这一监管状态转换，评价卖空策略业绩还面临另一个问题。即便没有涨点规则，许多股票，特别是小盘股或低流动性股，其实很难借到。比如你要卖空一只股票，你的经纪人就必须从其他人那里先借入这只股票（通常是大型共同基金或其他经纪客户），然后再借给你卖空。如果没人能够或者愿意借给你股票，你就无法卖空。因此，一个非常盈利的历史空头头寸，可能由于借不到股票而不可行。

上面所说的两次状态转换是较为明显且众所周知的。但是，还有其他一些与你的股票类别相关但非众所周知的细微状态转换，也会对策略业绩造成并不小的影响。我会在第 7 章中专门列出一节，探讨如何提出一个能自动侦测状态转换的模型。

5.5 小 结

自动交易系统是基于策略自动生成指令并将其传送给经纪账户的软件。这类软件的优点是：

- 保证与回测的策略完全一致。
- 不需手工操作，因此可以同时运行多个策略。
- 最重要的是，可迅速传送指令，这对高频交易策略尤其重要。

半自动交易系统和全自动交易系统的区别有：

- 半自动交易系统，交易员仍需向组合交易器或差价交易器手动上传包含指令详细信息的文本文件，并在恰当的时间手动点击按钮传送指

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

令。不过，指令文本文件可由 MATLAB 等程序自动生成。

• 全自动交易系统，程序能够在整个交易日（甚至多日）自动下载数据并传送指令。

自动交易系统完成后，你就要面对交易执行中的重要问题：最小化交易成本和仿真交易。最小化交易成本主要是别让你的指令规模过多偏离平均日交易量和市值。仿真交易可以：

• 发现策略和执行程序中的软件漏洞。
• 发现前视偏差或数据迁就偏差。
• 发现操作困难，进而规划操作流程。
• 更真实地估计交易成本。
• 获得有关盈亏波动、资本使用、投资组合规模和交易频率的重要直觉。

最后，在真实交易业绩不如回测业绩时该做些什么呢？通常先处理以下常见问题：消除策略或执行软件中的漏洞；降低交易成本；通过减少参数简化策略。但从根本上来说，你的策略可能还存在数据迁就偏差和状态转换影响。

如果你相信（并且只能相信，因为永远无法证明）真实交易业绩差只是因为运气不好，而不是由数据迁就偏差或状态转换造成，在面对继续交易和保全资本这两个截然不同的选项时，你该如何选择？这个关键问题将在下一章讨论，我们将系统讨论在面临亏损时如何保全资本并在市场好转时仍能东山再起的方法。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 6 章 资金和风险管理

所有策略都可能遭遇亏损，用专业术语来说就是挫跌。挫跌可能持续几分钟，也可能持续好几年。想从量化交易中赚钱，风险管理至关重要，把挫跌控制在可接受的范围内，把头寸建在净值的最优杠杆水平上，才能实现财富的最大可能增长。此外，如果不止一个策略，还需要在它们之间进行最优的资本配置，从而使风险调整后的整体收益最大。

6.1 最优资本配置和杠杆

假设计划进行几个策略的交易，每个策略都有其预期收益和标准差。那么，如何在这些策略之间进行最优的资本配置呢？此外，总杠杆（投资组合规模与账户净值的比率）又应该是多少呢？我在前言中提到的 Edward Thorp 博士的一篇论文中对这个问题进行了很好的说明（Thorp，1997)，本章将主要参考他的讨论。(Edward Thorp 博士的讨论关注证券组合，而我的讨论主要着眼于策略组合，不过所用的数学方法基本上是一样的。)

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

所有的优化问题都先要有一个优化目标。我们的优化目标是长期财富最大化，我相信这也是所有个人投资者的优化目标。长期财富最大化与投资组合的长期复合增长率 g 最大化是一回事。注意，这个目标意味着一定要避免赔光（净值为零或负数）。这是因为，如果在未来某时点赔光的概率大于零，则长期财富必然为零，长期增长率也肯定为零。

(在所有讨论中，假设所有的交易盈利均用于再投资，因此，杠杆复合增长率最为重要。)

假设策略 i（这里用 i 代表第 i 个策略）的收益率服从正态分布，其均值 m_i 和标准差 s_i 已给定（收益扣除融资成本，也就是超额收益）。这是金融中常用的假设，尽管可能并不十分精确。因为金融市场中较大亏损发生的概率会远远高于正态分布。然而，任何科学发明和工程试验都从最简单的模型、最粗略的假设开始，金融也不例外。本章后面的内容，我将讨论如何对此进行修正。

用列向量 $F^* = (f_1^*, f_2^*, \cdots, f_n^*)^T$ 表示分配到 n 个策略的最优净值比例，其中 T 代表转置。

给定优化目标并假设收益率服从正态分布，Thorp 博士给出了以下最优配置公式：

$$F^* = C^{-1}M$$

其中，C 表示协方差矩阵，矩阵的元素 C_{ij} 表示第 i 个策略和第 j 个策略收益率的协方差，-1 表示矩阵的逆，$M=(m_1, m_2, \cdots, m_n)$ T 表示策略平均收益率的列向量。注意，这些收益率是单期、单利（非复利）、无杠杆收益率。例如，如果策略是买入 1 美元的股票 A，卖出 1 美元的股票 B，一期后获利 0.10 美元，无论账户中的净值是多少，m 均为 0.05。

如果假设所有策略在统计上独立，协方差矩阵就变为对角矩阵，对角线元素等于每个策略收益率的方差，公式也十分简单：

$$f_i = m_i/s_i^2$$

这就是著名的凯利公式（围绕这个公式有许多有趣的故事，可参见

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

Poundstone，2005）在连续金融中的应用，它给出了一个特定交易策略的最优杠杆。

本章结尾提供了凯利公式在一个策略案例中的简单推导，有兴趣的读者可以看看。

例 6.1　一道有趣的智力题（风险为什么不是好东西）

这道智力题可能会难住不少职业交易员。假设某只股票的股价服从（几何）随机游走模型，也就是说每分钟以相同的概率上升或下降 1%。如果购买了这只股票，是最有可能盈利、亏损还是持平（长期持有并忽略融资成本）?

绝大多数交易员会脱口而出“持平!”，但实际上是错误的。正确的答案是，以每分钟 0.005%（0.5 个基点）的速度亏损！这是因为，对几何随机游走模型来说，平均复合收益率不再是短期（或单期）收益率 m（此处为零），而是 $g=m-s^2/2$。这与本章附录中给出的复合增长率 $g(f)$ 公式相同，其中杠杆 f 设为 1，无风险利率设为 0。这也与几何平均值通常小于算术平均值的事实相一致。（除非所有数字完全相同，否则这两个均值不会相等。）若我们假定收益率的算术均值为零，就像我所假定的那样，几何平均值（平均复合收益率）就必然是负的。

此处给我们的教训是，风险总是会减少长期增长率——由此可见风险管理的重要性！

* 这个例子是我的博客文《最大化复合收益率》的重新修正，原文见 epchan. blogspot. com/2006/10/maximizing-compounded-rate-of-return. html。

通常，由于参数估计存在误差，加上收益率不一定完全服从正态分布，交易员出于对安全的考虑，所使用的杠杆只有最优杠杆的一半。这就是“半凯利”投机。

如果你是零售交易账户，最大杠杆 l 只能为 2 或 4，究竟为何值取决

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

于你是持有当日头寸还是隔夜头寸。此时，必须通过乘以相同的因子 $l/(|f_1|+|f_2|+\dots|f_n|)$ 来减少每个 f_i，其中 $(|f_1|+|f_2|+\dots|f_n|)$ 是投资组合不受限制下的总杠杆。这里忽略了单个策略持有的头寸会相互抵消的情况（如一个多头头寸和一个空头头寸分别与相应的国债空头头寸和国债多头头寸相互抵消），头寸抵消可以使用比公式计算结果更高的杠杆。

采取上述的资本配置和杠杆，将使净值的长期复合增长率最大。那么最大复合增长率是多少呢？应为：

$$g=r+S^2/2$$

其中，S 即为投资组合的夏普比率。第 2 章提到，如果使用凯利公式给出的最优杠杆，投资组合（或策略）的夏普比率越高，净值（或财富）的最大增长率越高。下面的简单数学计算说明了这个结论。

例 6.2　用凯利公式计算最优杠杆

让我们来看一个使用凯利公式的例子。假设投资组合只包含跟踪标准普尔 500 指数的 ETF 基金 SPY 的多头头寸。设 SPY 的平均年收益率为 11.23%，年标准差为 16.91%，无风险利率为 4%。因此，投资组合的年平均超额收益率为 7.231%，已知年标准差 16.91%，可算出夏普比率为 0.4257。根据凯利公式，最优杠杆 $f=0.07231/0.1691^2=2.528$。（注意一个有趣的地方：凯利公式的 f 与时间长短无关，所以并不涉及是否年化收益率和标准差，这跟夏普比率与时间长短有关正好相反。）最后，考虑了融资成本的年化复合杠杆收益率为 13.14%。

你可以从雅虎财经下载 SPY 每日价格数据来验证这些结果，并可以在电子表格上计算不同数据。我验证 2007 年 12 月 29 日那一天，电子表格可从 epchan.com/book/example6_2.xls 下载。在 H 列，计算了 SPY（调整后）收盘价的日收益率，从 H 列 3 760 行开始，计算了 SPY 的（年化）平均收益率和标准差、投资组合的平均超额收益率和夏普比率、凯利

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

杠杆以及复合增长率。

这个策略的凯利杠杆 2.528 意味着，如果投资 10 000 美元现金，并且相信收益率和标准差的期望价值，就可借钱买入价值 252 800 美元的 SPY。此时，这 10 000 美元投资的预期年复合增长率为 13.14%。

作为比较，我们来算一下没有杠杆时的复合增长率是多少（见本章附录中的公式）：$g=r+m-s^2/2=0.1123-0.16912^2/2=9.8\%$。这是仅用现金购买 SPY 的长期增长率，不等于 11.23%的年收益率。

例 6.3　用凯利公式计算最优配置

我们选取三只特定行业的 ETF，来看看在它们之间如何进行资本配置，从而获得投资组合的最大增长率。这三只 ETF 是：OIH（原油服务）、RKH（区域银行）和 RTH（零售）。每日价格可从雅虎财经下载，在 epchan.com/book 上保存的文件名为 OIH.xls、RKH.xls 和 RTH.xls。以下是检索这些文件并计算 M、C 和 F* 的 MATLAB 程序（epchan.com/book/example6_3.m）。

```
% 清除以前定义的变量
clear;
% 文件 " OIH.xls" 读入 MATLAB 中
[num1, txt1] =xlsread ('OIH');
% 第一列（从第二行开始）是交易日期，格式为 mm/dd/yyyy
tday1=txt1 (2: end, 1);
% 将日期格式转化为 yyyymmdd.
tday1=datestr (datenum (tday1,'mm/dd/yyyy'),'yyyymmdd');
% 将日期字符串先转化为单元格数组，然后转化为数值格式
tday1=str2double (cellstr (tday1));
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 最后一列为调整后的收盘价格
adjcls1=num1 (:, end);
% 将文件" RKH.xls" 读取到 MATLAB 中
[num2, txt2] =xlsread (' RKH');
% 第一列 (从第二行开始) 是交易日期，格式为 mm/dd/yyyy
tday2=txt2 (2: end, 1);
% 将日期格式转化为 yyyymmdd.
tday2=...
estr (datenum (tday2,'mm/dd/yyyy'), 'yyyymmdd');

% 将日期字符串先转化为单元格数组，然后转化为数值格式
tday2=str2double (cellstr (tday2));
adjcls2=num2 (:, end);

%将文件" RKH.xls" 读取到 MATLAB 中
[num3, txt3] =xlsread ('RTH');
% 第一列 (从第二行开始) 是交易日期，格式为 mm/dd/yyyy
tday3=txt3 (2: end, 1);
% 将日期格式转化为 yyyymmdd.
tday3=...
datestr (datenum (tday3,'mm/dd/yyyy'),'yyyymmdd');

% 将日期字符串先转化为单元格数组，然后转化为数值格式
tday3=str2double (cellstr (tday3));
adjcls3=num3 (:, end);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 合并数据
tday=union (tday1, tday2);
tday=union (tday, tday3);
adjcls=NaN (length (tday), 3);

[foo idx1 idx] =intersect (tday1, tday);
adjcls (idx, 1) =adjcls1 (idx1);
[foo idx2 idx] =intersect (tday2, tday);
adjcls (idx, 2) =adjcls2 (idx2);
[foo idx3 idx] =intersect (tday3, tday);
adjcls (idx, 3) =adjcls3 (idx3);

ret= (adjcls-lag1 (adjcls)) ./lag1 (adjcls); % 收益率

% 找到收益率值缺失的日期
baddata=find (any (~isfinite (ret), 2));
% 去掉收益率值缺失的日期
ret (baddata,:) = [];
% 超额收益率：假设年无风险利率 4%
excessRet=ret-repmat (0.04/252, size (ret));
% 年平均超额收益率
M=252 * mean (excessRet, 1) '
% M =
%
0.1396
%       0.0294
%       -0.0073
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
C=252 * cov (excessRet) % 年协方差矩阵
% C =
%
%      0.1109 0.0200 0.0183
%      0.0200 0.0372 0.0269
%      0.0183 0.0269 0.0420
F=inv (C)  * M    % 凯利最优杠杆

% F =
%
%      1.2919
%      1.1723
%      -1.4882
```

注意，RTH 的平均超额收益率是负的，因此，凯利公式给出卖空建议也就不足为奇了。

你可能想知道使用这个最优配置所产生的夏普比率和最大复合增长率是多少。多策略高斯过程所对应的最大复合增长率为：

$g(F^*) = r + F^{*T} C F^* / 2$

夏普比率为：

$S = \sqrt{F^{*T} C F^*}$

下面是计算这两个量的 MATLAB 程序：

```
%最大年化复合增长率
g=0.04+F' * C * F/2
% g =
%      0.1529
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
S=sqrt (F' * C * F)
% 投资组合的夏普比率
% S =
%     0.4751
```

注意，投资组合的复合增长率为 15.29%，超过任一个单股的最大复合增长率。（经过验证你会发现，单期收益率最高的 OIH，其复合增长率为 12.78%。）

按照凯利公式，随着净值变化，需要对资本配置进行不断调整，以保证其最优。假如例 6.2 中，你按照凯利公式买入价值 252 800 美元的投资组合。第二天在 SPY 上亏损 10%，则投资组合的价值变为 227 520 美元，你的净值只有 74 720 美元。这时候该怎么办？凯利准则要求立刻将投资组合价值减少到 188 892 美元，因为这是最优杠杆 2.528 与净值 74 720 美元的乘积。

应至少在每个交易日结束时做一次资本配置的调整。除了更新资本配置，还应该周期性地跟踪计算最新的均值和标准差，从而更新 F^*。那么回溯期应为多久？更新凯利公式中这些输入变量的频率又是怎样的呢？这取决于策略的平均持有期。如果仅持有一天左右，根据经验规则，建议回溯期为六个月。使用较短回溯期的一个好处是，可以逐步减少正在失效策略上的风险暴露。至于更新频率，如果编写了一个程序，F^* 每天更新一次应该不是什么难事。

最后一点，有些策略每天生成的交易信号是一个变数，从而导致每天的头寸以及总资本也是变数。在这种情况下，该怎样使用凯利公式来决定资本呢？可以用凯利公式来计算最大头寸数量以及允许的最大资本。使用低于凯利公式计算出的杠杆永远是安全的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

6.2 风险管理

上一节讲到，凯利公式不仅可用于资本的最优配置及确定最优杠杆，还可用于进行风险管理。事实上，例 6.2 说明，在面临交易亏损时，凯利公式会建议你减小投资组合规模。无论风险管理方案是否基于凯利公式做出，亏损减仓都是风险管理的惯常做法。

无论何时发生亏损，风险管理总会建议要求减仓。（相反，当策略盈利时，最优杠杆会建议增仓。）许多研究员都持这种减仓态度，因此，当一家大型对冲基金面临重大亏损时，“金融传染”会影响许多其他大型对冲基金。

一个例子是 2007 年夏天的大崩盘，Andrew Lo. 和 Amir Khandani 的文章《2007 年 8 月，宽客到底怎么了?》对此进行了描述。在 2007 年 8 月，在住房抵押贷款违约风险的阴云笼罩下，许多知名对冲基金遭受了前所未有的亏损，如高盛的全球阿尔法基金亏损了 22.5%。数十亿美元资产在一周之内蒸发殆尽。即便是文艺复兴科技公司，这家一直以来被认为是最成功的量化对冲基金，在 8 月的上半月也亏损了 8.7%，不过后半月基本弥补上了亏损。除了亏损金额之大令人震惊，分布面之广也引起了金融界的极度关注。最奇怪的是，几乎没有基金持有住房抵押贷款支持证券，从表面上看这是恐慌的根源。因此，这成为对冲基金引发“金融传染”的典型案例。

这种传染发生的原因是，若一家对冲基金遭受大额亏损，它会卖出所持有的大额头寸（不论这些头寸是否就是一开始发生亏损的头寸）。抛售导致了证券价格的下跌（若是空头头寸则是上升）。如果其他对冲基金也持有类似头寸，很快也会遭受亏损，风险管理系统要求其也抛售头寸。如此循环。例如，2007 年夏天，一家大型对冲基金持有次级住房抵押贷款支持证

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

券，并遭受了巨额亏损。风险管理系统要求抛售投资组合中流通股票头寸，但在那时这些头寸可能尚未受到次贷崩溃的影响。由于股票头寸的抛售，其他未持有住房抵押贷款支持证券的统计套利对冲基金现在也可能遭受大额亏损，进而也卖出它们的股票。因此，住房抵押贷款支持证券市场的抛售，突然演变成股票市场的抛售，这就是对"传染"一词的完美解释。

为了按照凯利公式进行操作，必要时要承担实际亏损，并且要不断调整投资组合的规模和交易频率。因此，绝大多数交易员更愿意采用半凯利杠杆进行交易就很容易理解了。较低的杠杆意味着，风险管理所要求的减仓量也较小。

有时，采取保守的半凯利杠杆仍然显得过于激进，交易员可通过其他约束来进一步限制投资组合的规模。因为正如前面所说，凯利公式适用于连续金融的前提是收益率呈正态分布。（从感觉上看金融是连续的，因为金融市场中产生的盈利和亏损是一个连续变化的过程，这与纸牌游戏中的离散损益不同。）但收益率分布未必完全与正态分布吻合，因为发生巨额亏损的概率要远远高于正态分布完美的钟形曲线所给出的可能值。这种收益率的真实分布有时被称作"厚尾"，意为事件发生概率极大偏离均值的可能性要远远高于正态分布钟形曲线所给出的情况。Nassim Taleb 将这种发生概率极低的事件称为"黑天鹅"事件（见 Taleb，2007）。

可通过对历史最大单期亏损的简单回测，来估计落在正态分布之外的极端事件。（单期可以是一周、一天或一小时。唯一的准则是，在每个单期末，要用凯利公式对投资组合进行调整。）对于你能承受的净值最大单期挫跌要做到心中有数。最大可承受单期挫跌与最大历史亏损的比较结果，会告诉你半凯利杠杆是否仍然过大。所使用的杠杆，一般是半凯利杠杆和在最大历史亏损下所能获得的最大杠杆中较小的一个。1987 年 10 月 19 日，著名的"黑色星期一"，标准普尔 500 指数出现了最大历史单日亏损 20.47%。如果你仅能承受一天 20%的净值亏损，可用的最大杠杆就是 1，而半凯利杠杆高达 1.26。因此，在这种情况下，即使是半凯利杠杆也

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

不够保守，不足以在“黑色星期一”中幸存下来。

风险管理中，最可怕的情形是历史上从未出现过的情形。用路德维格·维特根斯坦的话说：“凡是不能说的事情，就应该沉默。”对这种不可知，理论模型应该沉默。

止损是风险管理的好方法吗？

有些交易员认为，好的风险管理就是对每笔交易采取止损，即当头寸亏损达到一定的百分比时，交易员就要清仓。采取止损可以避免投资组合出现灾难性巨额亏损，是一个普遍的误解。当灾难发生时，证券价格是不连续下跌的，因此，清仓止损指令的成交价格要远低于灾难发生前的价格。所以，在清仓行为中，止损不是避免灾难性亏损，而是使其真实发生。止损只有在惯性（趋势）状态时，才是有益的。换句话说，在交易的期望持续时间内，价格会进一步下跌。否则，如果市场在这段时间内是均值回归的，不过快的清仓才能最终弥补亏损。

当然，要判断市场是处于惯性状态（止损有益）还是均值回归状态（止损有害），并不是一件容易的事。我的观察是，若消息或其他基本面原因（如公司业绩下降）导致价格波动，一般是惯性状态，用交易员的话说，这时“不要试图与迎面而来的火车相撞”。例如，如果对一家公司进行基本面分析表明其价值被高估，股价就很可能要逐渐下跌（起码相对市场指数）到一个新的、较低的均衡价格。只要公司的基本面不发生改变，这种下跌就是不可逆转的。然而，当证券价格在没有任何消息或明显原因的情况下剧烈波动时，很有可能是流动性事件引起的，如大股东因个人特殊原因大量抛售，或大投机者突然停止做空。流动性事件的持续时间较短，价格也很可能回归到之前的水平。

第 7 章将对适合均值回归策略和惯性策略的清仓策略进行更详细分析。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

除了头寸风险（包括市场风险和特有风险），还存在其他形式的风险，按发生可能性从大到小，有模型风险、软件风险和自然灾害风险。

模型风险是指由于交易模型错误而非市场统计特征的变化而造成交易亏损的可能性。错误发生的原因可能有很多，第 3 章已经提到过一些：数据迁就偏差、存活偏差等。为消除回测程序中各种偏差和错误，请合作伙伴或咨询顾问独立地复制出你的回测结果来确保其有效性是很有帮助的。结果复制是科学研究中的常用做法，在金融研究中也是必不可少的。

模型风险也可能来自于使用相同策略的机构交易员的激烈竞争，而不是由模型或回测程序的偏差或错误造成。也有可能是因为市场结构基本特征变化削弱了交易模型的优势，即第 5 章讨论过的状态转换。

随着交易亏损的增大，除了逐渐降低模型的杠杆直到零，没有什么更多的办法可以降低模型风险。根据最新的历史收益率均值和标准差，使用凯利公式不断调整杠杆，可以系统地做到这一点。（随着回溯期内的历史均值降为零，凯利杠杆也将调整为零。）这比突然弃用一个发生了很多挫跌的交易模型要好（参见下一节“做好心理准备”。）

软件风险是指由于每日生成指令的自动交易系统没有真实反映回测模型所引起的风险。其发生的原因在于软件故障无处不在。第 5 章讨论过消除软件错误的方法：比较自动交易系统生成的交易与回测系统生成的理论交易，看它们是否相同。

最后，像地震、海啸之类的物理或自然灾害的发生也会带来巨大损失。在建立对冲头寸之前网络连接中断了怎么办？当传送指令时断电了怎么办？避免物理灾害造成交易中断的许多方法，可参见第 4 章“设备”一节。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

6.3 做好心理准备

一本量化交易的书用一节来讲心理准备的内容，看起来似乎有点奇怪。毕竟，量化交易不正是要将我们从情绪中解放出来，让电脑按部就班地做出所有交易决策吗？事情并非如此简单，没有做好心理准备的交易员，尤其是在头寸或某日的盈亏出现异常时，往往会推翻自动交易系统的决策。因此，即使用量化策略进行交易，了解一些我们自己的心理弱点也是十分必要的。

幸运的是，专门有“行为金融”（Thaler，1994）研究非理性金融决策。我将概括介绍一些常见的影响交易的非理性行为。

第一种行为偏差被称为“禀赋效应”、“安于现状偏差”或“亏损厌恶”。前两个效应会使交易员长时间持有亏损头寸，因为交易员（和一般人）更偏好安于现状（安于现状偏差），或者只有在价格比购买时高很多的情况下才愿意卖掉股票（禀赋效应）。正如“资金和风险管理”一节谈到的，若持有亏损头寸要有合理原因（如期待均值回归）。然而，这些行为偏差会让交易员在没有任何合理原因的时候，也持有亏损头寸（如预期的市场趋势将使你的头寸亏损更多）。同时，亏损厌恶会导致一些交易员过早地卖掉他们的盈利头寸，尽管平均来说多持有一段时间会获利更多。为什么他们会早早地卖掉盈利头寸呢？因为失去手头盈利所带来的痛苦，远高于获得更高盈利所带来的快乐。

当某人错误地建立了头寸（由于软件故障、操作失误或数据问题等原因）并发生大额亏损时，行为偏差会最明显、最惨烈地表现出来。一旦发现错误，理性的做法是立刻清仓。然而，交易员往往幻想等待均值回归，从而能减少交易亏损。除非均值模型表明当前就是建仓的最佳时刻，等待均值回归很可能导致更大的亏损。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

另一个常见的偏差是“代表性偏差”，即人们倾向于对近期经验赋予过多权重，而低估了长期平均的作用（Ritter，2003）。大额亏损后，交易员，甚至是量化交易员，倾向于立刻修改策略中的某些参数，以避免这种大额亏损。如果他们使用修改后的系统进行交易，则这样的修改是不明智的，因为可能带来其他尚未发生的大额亏损，或者消除许多现有的盈利机会。我们必须牢记我们是在概率状态中进行操作：没有系统能避免所有会导致亏损的市场异常状况。

如果你认为系统真的有缺陷，并打算调整，同样要对修改后的版本进行回测，以确保它能够在足够的回测期、而不仅仅是过去几周内的业绩优于原系统。

交易员比经济学家更熟知两个心理弱点：恐惧和贪婪。

当交易模型处于巨大、持续的挫跌中时，会引起恐惧。许多交易员（以及他们的经理、投资者等）要承受彻底关闭模型的巨大压力。其他过度自信且轻率冒险的交易员会进行反向操作：他们会在亏损模型上加倍投注，寄希望于模型反弹并在反弹之时最终挽回损失。这两种行为都是不理性的：如果按照凯利公式管理资本配置和杠杆，应该逐步降低分配给亏损模型的资本。

当模型运行良好并获利颇丰时，贪婪是一种更常见的情绪。此时的冲动是迅速增加杠杆以期快速致富。而一个训练有素的量化交易员，仍会将杠杆保持在凯利公式给出的水平之下，同时对厚尾事件发生的可能性保持警觉。

恐惧和贪婪都会导致过度杠杆化：在恐惧时，人们会尝试通过追加资本来挽回损失；在贪婪时，人们会在策略刚开始盈利时，就过快追加资本。因此，风险管理的黄金法则是：任何时刻都要将投资组合的规模保持在可控范围内。然而，知易行难。一些著名的大型基金公司就因冲动而过度杠杆化，最终破产，如 2000 年的长期资产管理公司（Lowenstein，2000）和 2006 年的 Amaranth Advisors（epchan. blogspot. com/2006/10/

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

highly-improbable-event. html)。在 Amaranth Advisors 的案例中，同一交易员（Brian Hunter）在同一策略（天然气跨期套利交易）上采用了过大的杠杆，导致 60 亿美元的亏损，轻松赔光了基金的净值，简直是一个教科书式的案例。

我在机构工作和个人交易中，都曾经受过这样的压力，并且还都是在初期。我在一家资金管理公司工作时，给基金投资者造成了超过 1 亿美元的损失。因为出于一点点贪婪之心，我对一个投资组合追加了 1 亿美元的资本，而这个投资组合所依赖的策略才仅仅交易了 6 个月。（那时我还没学到凯利准则。）似乎是得到的教训不够深刻，我开始进行独立交易时又犯了同样的错误，是关于 XLE、一只能源交易型开放式指数基金（ETF）和原油期货（CL）的均值回归差价策略。当差价一直没有出现均值回归的趋势时，我很顽固地将策略的交易规模增加到 50 万美元。最后，当亏损高达 6 位数的时候，我因为恐惧对这个差价头寸清仓。自然，在我清仓后不久，差价开始回归了。（幸运的是，在我进行独立交易的第一年，其他几个策略的业绩很好，所以在第一个交易年份结束时，我只有轻微的亏损。）

我们该如何训练自己克服这些心理弱点，学会避免人为去否定模型的决策并迅速而正确地改正交易错误呢？如同大多数人的尝试一样，可以先从小额组合做起，逐渐做好心理准备、训练有素并获得对模型的信心。当你有足够的心理承受能力去应对每日盈亏波动，并且能够抑制心理上的原始冲动时，投资组合的实际业绩便将会趋近策略的理论预期业绩。

在经历了上面提到的那些灾难性交易之后，我才找到了这些应对它们的办法。而我在凯利公式中所获得的训练和信心进一步防止了类似事件的再次发生。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

6.4 小 结

风险管理是交易中至关重要的准则。由于单个交易或极短时间内的巨额亏损，导致大型对冲基金和投资银行倒闭的例子不少。实际上，大多数亏损是由于头寸的过度杠杆化而非模型本身的错误所造成的。一般来说，如果模型业绩不佳，交易员就不会过度使用杠杆。只有那些迄今为止业绩仍十分出色的交易模型，因过度自信和过度杠杆化，处于巨额亏损的最大风险之中。本章提供了一个重要的风险管理工具：用凯利公式决定最优杠杆。

除了决定最优杠杆，凯利公式还有另外一项用途：根据不同策略收益的协方差，决定不同策略之间的最优资本配置。

但是，如果对交易的盈亏波动没有做好心理准备，从而偏离理性决策的方案（比如你的交易模型），再好的风险管理公式或系统也不能避免灾难的发生。风险管理的终极理念其实十分简单：不要被恐惧和贪婪所左右。为了获得这样的心理素质，开始交易时必须投入少量资金慢慢做，在用凯利公式扩大交易规模之前，要对交易的各个方面进行彻底测试（模型、软件、操作流程、资金和风险管理）。

我发现，若想慢速、谨慎地发展，拥有其他的收入来源或业务有助于保持财务和情绪的稳定（避免由于缓慢发展而引起的烦躁）。确实应该寻找一件事情来分散注意力，不论它是否能够带来收入，都会有助于你财富的长期增长。

6.5 附录：收益率正态分布时凯利公式的简单推导

假设策略（或证券）的收益率呈正态分布，很容易就能推导出凯利公

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

式。适用于正态分布的复合杠杆增长率公式为：

$g(f)=r+fm-s^2f^2/2$

其中，f 为杠杆，r 为无风险利率，m 为平均非复合单期超额收益率，s 为非复合单期收益率的标准差。这个公式的推导也很简单，但不如凯利公式那么简单，留给读者练习，可参考前面提到的 Thorp 的论文。

为了得出使 g 最大化时的 f，令 g 对 f 的一阶导数为零：

$dg/df=m-s^2f=0$

由等式可得 $f=m/s^2$，即为正态分布下策略或证券的凯利公式。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 7 章 量化交易专题

本书前 6 章涵盖了研究、开发及执行量化策略所需的大部分基础知识。本章将对量化交易的重要主题进行详细讲解。这些主题构成了统计套利交易的基础，并为大多数量化交易员所熟悉。此外，这些主题对于形成交易直觉也非常有帮助。

我将首先介绍两类基本的交易策略：均值回归策略和惯性策略。均值回归时段和趋势行为时段，就是交易员所说的“状态”的例子，不同状态之间的转换正是本章要重点讨论的一个专题。均值回归策略的数学基础源于时间序列的平稳性和协整性概念，我在后面会讲到。然后，我将介绍一个许多对冲基金用于管理大额组合并曾导致其业绩大幅波动的理论：因子模型。交易员经常讨论的其他策略类型还包括季节性交易策略和高频策略。所有的策略都要有清仓方式，我将介绍各种清仓方式中所蕴含的不同的逻辑。最后，我将讨论如何才能最大程度提高策略的收益：使用更高的杠杆还是交易高贝塔股票？

7.1 均值回归策略和惯性策略

只有当证券价格是均值回归的或趋势的，交易策略才能盈利。否则，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

价格是随机漫步的，交易将无利可图。如果你相信价格是均值回归的，并且目前相对较低，应当现在买入，并准备在以后价格升高时卖出。但是，如果你相信价格是趋势的，且目前处于低位，应当现在卖出（卖空），并准备在以后价格更低时买入。价格处于高位则刚好相反。

学术研究表明，股票价格“一般而言”非常接近随机漫步。但这并不意味着在特殊条件下价格不会表现出一定程度的均值回归或趋势行为。此外，根据选取的时间段不同，同一时刻的价格既可以是均值回归的，也可以是趋势的。因此，构建一个策略，实际上就是要判断，在特定条件和特定时间段，价格究竟是均值回归的还是趋势的，以及任何给定时点的初始参考价格是多少。（当价格是趋势的，就称其具有“惯性”，因此，相应的策略常被称为“惯性策略”。）

价格在同一时间既均值回归又趋势的现象，被有些人描述为股票价格具有“分形”性质。技术分析者或图表分析者喜欢用所谓的艾略特波浪理论来分析这种现象。还有人喜欢用机器学习或人工智能（特别是像隐马尔科夫模型、卡曼滤波、神经网络等技术）来发现价格是否处于均值回归或趋势“状态”。我个人还没发现这些均值回归或惯性的理论特别有用。（不过，在关于状态转换的那一节，提供了一个成功预测某只股票状态转换的例子。）我倒是发现这样假设通常是安全的：除非公司的预期盈利发生了变化，股票价格会均值回归。实际上，金融学者（Khandani and Lo, 2007）已经建立了一个非常简单并且能够持续盈利（不考虑交易成本）的短期均值回归模型。当然，均值回归是否够强够长，以保证在引入交易成本后仍有利可图，则是另外一回事了。而找到均值回归够强够长的特殊情形，正是交易员的任务。

尽管均值回归非常普遍，回测一个盈利的均值回归策略可能很不可靠。

许多历史金融数据库中都包含报价错误，而这些错误常常会人为抬高均值回归策略的业绩。原因很简单：均值回归策略买入时用到一个比某一

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

移动平均低的、虚构的报价，却按下一个与移动平均一致的、正确的报价卖出，从而获得盈利。只有在彻底清除了虚构报价之后，才能够完全相信均值回归策略的回测业绩。

正如第 3 章所讨论的，存活偏差也会影响均值回归策略的回测。经历极端价格的股票很可能被收购（股价涨得非常高）或破产（股价变为零）。均值回归策略会卖空前者、买入后者，而这两种情况都会亏钱。而如果有存活偏差，这些股票都不会出现在历史数据库中，从而导致人为抬高回测结果。你可以从表 3－1 中知道哪些数据库有存活偏差。

惯性可能产生于信息的缓慢扩散——当越来越多的人开始注意到某条新闻，越来越多的人决定买入或卖出某只股票，就使得价格朝某一方向运动。前面说过，当公司的预期盈利变化时，股票就会显示惯性。这可能发生在公司公布其季度盈利之时，投资者逐渐了解该公告，或通过多步执行大额指令以应对该变化（使市场冲击最小）。事实上，根据这种现象，我们可以构建一个叫做“后盈利公告漂移”（PEAD）的惯性策略。（更多关于这个策略的有用参考文章，可参见 quantlogic. blogspot. com/ 2006/ 03/ pocket-phd-post-earning-announcment. html。）实质上，这一策略建议你在盈利超出预期时买入股票，低于预期时卖空股票。更一般的，很多信息公告都有改变股票未来的预期盈利的潜力，从而又激发一个趋势时段。至于何种消息会触发趋势时段，以及趋势时段会持续多久，就要再一次靠交易员自己去发现了。

除了信息的缓慢扩散，大额指令因流动性需求分拆执行或大投资者的私募投资决策，也会导致惯性。相比于其他原因，这可能导致更多的短期惯性。随着大型经纪商越来越多地使用日益复杂的执行算法，要判断一个观察到的趋势背后是否存在一个大额指令变得越来越难。

惯性也可能由投资者的羊群行为所引发的：投资者把其他人的（可能是随机且无意义的）买卖决策作为自己交易决策的唯一判断标准。就像耶鲁大学的经济学家 Robert Schiller 在《纽约时报》（Schiller，2008）上说

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

的，没有人拥有他需要的所有信息以便做出一个完全有根据的财务决策。人们不得不依赖其他人的判断。然而，我们没有办法分辨他人判断的优劣。更成问题的是，人们是在不同的时点做出财务决策，并非在市政大厅一起开会时达成共识。第一个支付高房价的人在“告知”其他人房子是一项好投资，使得另外一个人也做出相同的决定，进而越来越多。因此，第一个购房者原本可能错误的购房决定，就会被当成“信息”扩散给其他人群。

不幸的是，这两种原因（个人流动性需求和羊群行为）引发的惯性状态的时间周期是极难预测的。如何才能知道某机构需分步执行的指令有多大？如何预测“羊群”效应何时会足够大到能形成一边倒的局面？不可捉摸的转折点在哪里？如果没有估计时间周期的可靠方法，我们就无法通过惯性交易获利。下一节“状态转换”将考察试图预测这些转折点或“拐点”的方法。

比较均值回归策略和惯性策略，还有最后一点值得考虑。使用相同策略的交易员之间日益激烈的竞争会对策略本身造成怎样的影响？对于均值回归策略而言，典型的结果就是套利机会的逐步消失，从而使得收益率逐渐降低至零。当套利机会消失殆尽时，均值回归策略就危险了，因为越来越多的交易信号来自于股票估值的基本面变化，而这并不会均值回归。对于惯性策略而言，竞争的后果是减少该趋势持续下去的时间周期。当消息以更快的速度扩散，从而更多交易员们更早地利用了这一趋势，均衡价格就会更快实现。所有在均衡价格实现之后建仓的交易员都将无利可图。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.2 状态转换[①]

状态是金融市场中一个最基本的概念。如果没有状态，何来“牛市”和“熊市”？自金融市场诞生之日起，人们就试图预测状态转换，即寻找所谓的“拐点”。

哪怕我们能够稍微成功地预测牛市和熊市的转换，就只要讨论这种转换即可，不用讨论其他的。但事实上，这种预测并不容易。也正因为预测这种转换十分困难，使得很多研究者在金融市场上寻找其他状态转换，希望能找到更适合现有统计工具的状态转换。

我已经介绍过由于市场和监管结构改变而引起的两种状态转换（这两个例子中，状态没有再转换回之前的状态）：2003 年股票价格改为十进制和 2007 年初涨点规则的废除（具体见第 5 章）。这些状态转换由政府政策变化所导致，所以我们并不关注于这种类型的变化。当然，能准确预测政策变化的人也寥寥无几。

其他最常见的金融或经济状态研究，包括通货膨胀与经济衰退状态、高波动率与低波动率状态以及均值回归与趋势状态。其中，波动率状态转换似乎最适用经典计量经济学工具，如广义自回归条件异方差（GARCH）模型（见 Klaassen，2002）。这并不奇怪，因为金融经济学家远早于股票价格本身而对波动率成功建模。然而，对于期权交易员很有价值的波动率状态转换预测，对股票交易员却没有帮助。

学术界一般沿用以下思路对股票价格的状态转换进行建模：

1. 假设价格在两个（或多个）状态上的概率分布不同。最简单的情

① 本节改编自我发表在 *Automated Trader* 杂志上的一篇文章。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

况，两个状态的价格都服从对数正态分布，但均值和（或）标准差不同。

2. 假设状态之间存在某种转移概率。

3. 使用诸如最大似然估计这样的标准统计方法，通过拟合历史数据，来确定状态概率分布和转移概率的参数。

4. 根据上述拟合模型，找出下一个时间步长的期望状态，更重要的是，找出股票的期望价格。

这种方法通常被称为“马尔科夫状态转换模型”或“隐马尔科夫模型”，这一模型通常基于贝叶斯概率框架。对这种方法感兴趣的读者，可以参见 Nielsen and Olesen（2000）、van Norden and Schaller（1993）或 Kaufmann and Scheicher（1996）。

虽然理论框架完美，马尔科夫状态转换模型在实际交易中的用途却不大。这是因为模型假设状态之间的转移概率都是固定的。实际应用中，这意味着在任何时候（正如 Nielsen 和 Olesen 的论文中所阐释的那样）股票从正常的静止状态转移到不稳定状态的概率非常小。而这对于想知道转移概率在何时（以及何种情况下）突然达到峰值的交易员，是完全没有用的。这就需要用到拐点模型。

拐点模型使用了数据挖掘方法（Chai，2007）：输入所有可能预测拐点或状态转换的变量。变量包括当前的波动率、最近一期收益，以及消费者信心指数、石油价格变化、债券价格变化等宏观经济数据的变化等。经济学家 Robert Schiller（2007）关于房地产市场拐点的文章受到广泛关注，文章指出：媒体上开始日益频繁地讨论繁荣或萧条实际上是下一个拐点的预测。

例 7.1 介绍了应当如何通过数据挖掘来确定拐点，这种方法建立在简单的技术指标之上，即股票价格序列作为输入变量、股票在多个持有期的收益率作为输出变量。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例 7.1　使用机器学习工具在股市的状态转换中获利

正如正文所讨论的，我相信数据挖掘方法能很容易地确定状态转换：检验大量的指标来确定究竟哪个能够预测转换。这是一个即使用 MATLAB 工作量也非常大的任务。但幸运的是，最近出现的一个机器学习程序，能使这个任务在几个小时之内完成。

我使用的工具叫 Alphacet Discovery，一个由 Alphacet 公司（www.alphacet.com，也是我公司的客户）新开发的集回测和执行于一体的平台。这一平台不仅整合了策略原型、回测、分析和实时部署所需的所有历史和实时数据，还包括适用于数据挖掘的机器学习程序，如神经网络和遗传算法。

我选择一只出色经纪商的股票 GS 作为金融行业的代表。我的目标是能否发现这个行业牛市熊市的拐点。初始假设是：利率重大变化、政府宏观数据发布或盈利公告会触发拐点。在写作本书的时候，Alphacet 还未将宏观数据和公司数据整合进数据库，所以我用 GS 股价的较大变化代表这类消息的发布。另外，无论何时，当 GS 股价在此大跌或大升之前达到其 N 天内的最高价或最低价，都很好地预示着前一状态快要结束了。所以我把这一情况也作为另一输入变量。

我们面临的问题是：多大的价格变动才足以触发状态转换？N 应该取多少天？新的状态通常会持续多久？（换句话说，最优持有期多长？）如果用老式的人工方法来回答这些问题是非常耗时的，因为必须在对自变量用多个阈值和对因变量用多个返回形态的情况下进行多次模拟。让我们看看 Alphacet Discovery 是怎样帮助我们实现流程自动化的。

模型的自变量是 GS 的日收益率，因变量是 GS 在不同持有期的未来收益率。Discovery 能很容易地找到得出最佳回测业绩的一个最优规则，或几个最优规则的集合。在我们的例子中，每一个百分比变化的阈值都可以被封装成一个规则。我给买入和卖出各设定两个阈值：−1%、−3%、

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

1%、3%。同样地，每一个持有期也可以被封装成一个规则，我设定六个持有期：1 天、5 天、10 天、20 天、40 天和 60 天。

准备此次研究所需的价格、百分比变化和 10 日内高/低时间序列非常容易，在 Discovery 里，大部分任务通过移动鼠标便能完成。（为了简便，我选择了 N 为 10，但这一参数也可优化。）我将一个 GS 的价格序列拖入编辑器并确定 1 日的频率。（价格序列开始于 2006 年 11 月。）见图 7.1 框 S1。然后，我向编辑器拖入预先设定好的能够计算 1 日百分比变动的“规则”和一个简单的 10 日移动高和低序列（参见图 7—1 框 I2）。接着，我通过用箭头将符号组框连向程序组框的方式就将最初的价格序列输入到程序组框。

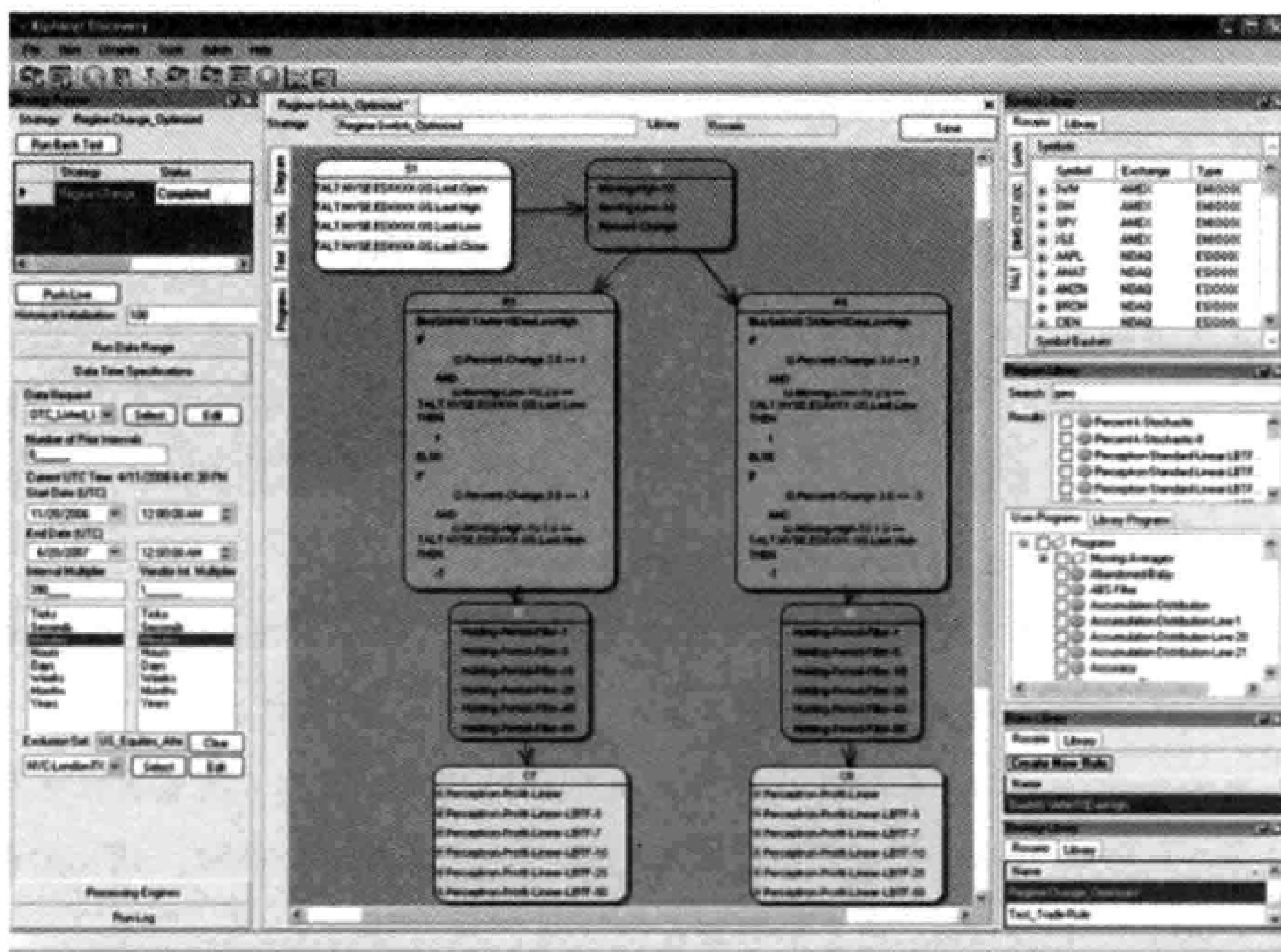


图 7—1　策略编辑器

现在，我可以通过使用下拉菜单创建条目并在一个新的规则编辑器进行输入。图 7—2 显示了基于±1%变化的买卖规则下规则框 R3 的变化。我为±3%也创建了一个类似的规则框。值得注意的是，在默认情况下，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

后续输入将会覆盖之前的输入。

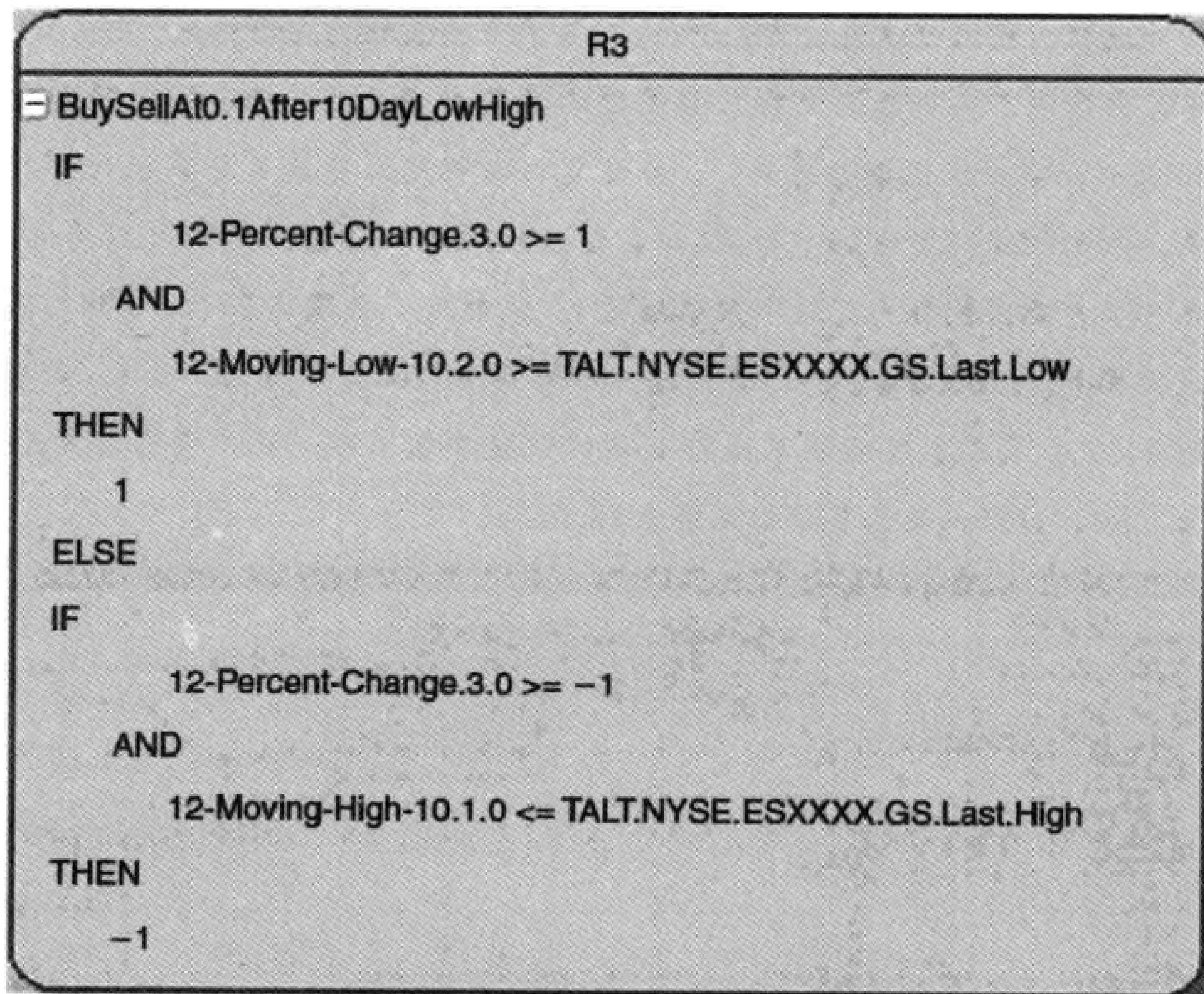


图 7—2 规则框

我们可以在预装的“持有期”程序中用不同参数来指定持有期。（实际上，如果你会编程，也可以自己创建相应的程序。）这些都封装在框 I5 中（±2%的算法单独列在 I6 中）。我们通过连接箭头的方式将框 R3 的结果输入到 I5 中，同样，将框 R4 的结果输入到 I9 中。

最后，我们对 I7 和 I9 的结果运行感知器学习算法（感知器是神经网络的一种）。基于最大化总盈利为目标的历史训练数据移动窗口，这种算法可以得出不同持有期下不同规则的最优权重。基于这些最优权重，感知器会在每个期限末触发买卖决策。

有趣的是，即使移动窗口的规则是以 N 天构建的，感知器也并不会强制我们一定处在 N 天的状态。每日，程序会根据移动窗口中的最新数

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

据和不同规则的线性权重结果得出最新参数优化，进而做出买卖决策。

现在，我们可以通过 Discovery 的图表来看策略的业绩结果。图 7—3 中，我展现了感知器优化后三条最佳的净值曲线。最佳的一条曲线是由 50 天移动窗口优化模型得到。（移动窗口的长度本身就是可以优化的对象，但此处我们省略这个步骤。）在图表应用的侧边栏上我们可以看到，在 6 个月的回测期内，这个策略通过 89 次往返交易获得 37.93%的总累计收益率。（相比较而言，GS 的买入持有策略的收益率为 15.77%，挫跌为 14%。）我们也展示了不同持有期的最佳净值曲线（I5：持有期为 10 天，R3：1%），其总累计收益率为 18.55%。

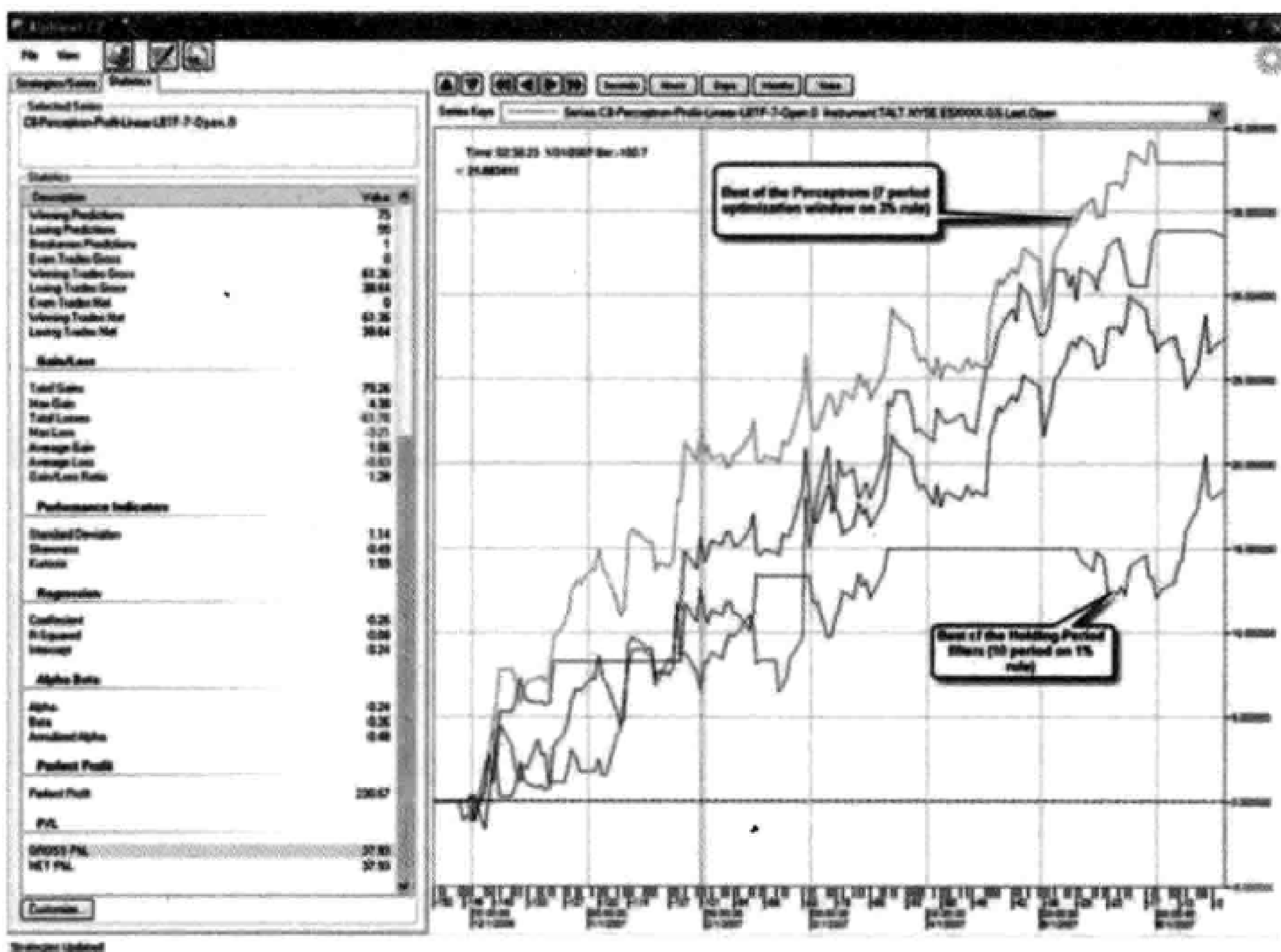


图 7—3　展示净值曲线的图表和展示业绩数据的侧边栏

虽然回测期很短，但收益率却令人印象深刻。是什么地方出错了吗？特别是，已经蔓延到所有基于机器学习或人工智能的策略的数据迁就偏差存在吗？Alphacet Discovery 的基本原理就是避免这种偏差的出现。尽管在这里我们没有特别说明，但从理论上来说，后顾移动窗口中的所有算法

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

和参数都可以优化，所以我们绝对没有使用预测不到的未来数据用于回测。当然，数据迁就偏差还是会出现，因为当回测业绩较差时我们会舍弃一类模型，并且尝试另一类新的模型，直到其回测业绩良好为止。但那个时候，我们已经处于回测业务，是不可避免的。

在这种情况下，还应注意搜索引擎优化的参数实际上相当有限：仅仅是持有期不同而已。这也进一步降低了数据迁就偏差风险。

因为回测看起来不错，我可以立刻按下一个按钮，对实时数据在仿真账户或实时交易账户中生成指令。

正如你所看到的，要创建一个状态转换模型并不是很难，只需用最简单的技术指标，以及能够有效地优化大量参数，并能在后顾移动窗口中被严格执行。（见第 3 章的无参数交易模型。）如果我们可以确认价格移动与宏观经济及公司信息之间的关系，结果可能会更好。我相信这种技术会给许多交易所交易基金、期货甚至现货交易带来盈利。

7.3 平稳性和协整性

如果一个时间序列不会越来越大地偏离初始值，这个时间序列就是“平稳的”。用专业术语来说，平稳的时间序列就是“零阶自积”的，即Ⅰ(0)。（见 Alexander，2001。）显然，如果证券的价格序列是平稳的，它就很可能适用于均值回归策略。不过，大多数股票的价格序列都不是平稳的，通常表现为几何随机游走，不断地离初始点（如首次公开发行）价值越来越远。尽管如此，你能找到像买入一只股票、卖出一只股票这样的股票配对，配对的市场价值是平稳的。这种情况下，两个独立的时间序列被称为“协整”。之所以这样称呼，是因为两个时间序列是零阶自积的。通常，协整配对中的两个股票来自同一行业。交易员早已很熟悉这种配对

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

交易策略。他们在配对的差价低的时候买入配对组合，在差价高的时候卖出（卖空）配对——这就是经典的均值回归策略。

一个协整价格序列配对的例子是例 3.6 的黄金 ETF（GLD）与采金企业 ETF（GDX）。构建一个组合：一份 GLD 多头和 1.6766 份 GDX 空头，组合的价格就是平稳的时间序列（见图 7—4）。GLD 和 GDX 的具体份数是通过对两个价格序列回归得到的（见例 7.2）。

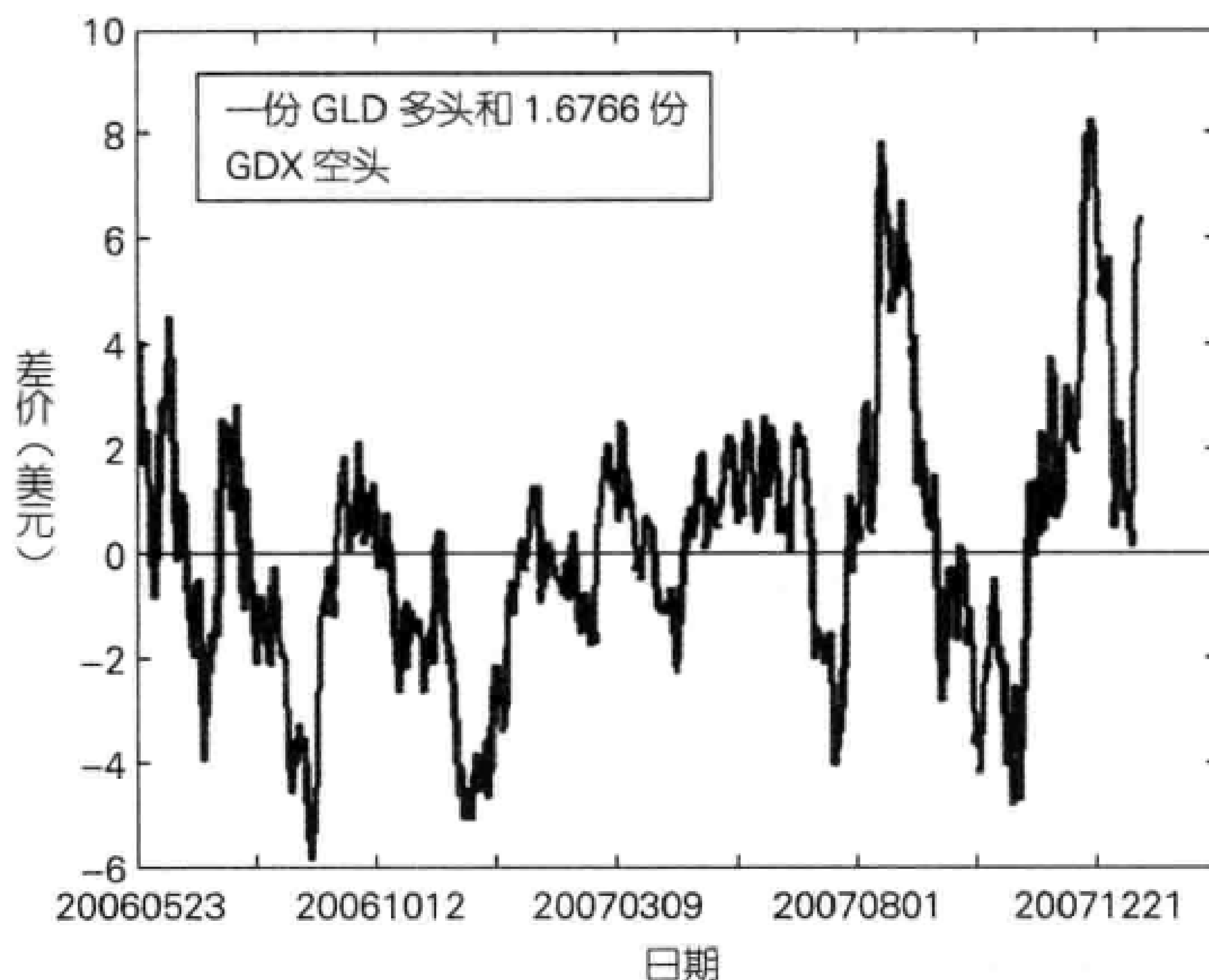


图 7—4　GLD 和 GDX 差价形成的平稳时间序列

例 7.2　如何构造协整度高（均值回归）的股票配对

正文说过，如果做多一个证券的同时以正确的比例做空另一个同行业证券，组合（或“差价”）有时是一个平稳序列。平稳的时间序列是均值回归策略的绝佳候选。本例将教你如何从 www.spatial-econometrics.com 免费下载 MATLAB 工具包来检验两个价格序列是否协整，以及如何找到最优的对冲比率（即第二个证券份数与第一个证券份数的比例）。

协整检验的主要方法是 ADF 检验，对应的函数名为 cadf。对这一方法的详细描述也可在上面说的网站上找到。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

下面的程序同样可以从 epchan.com/book/example7_2.m 下载：

```
% 清除空间已有变量
clear;
% 读入 " GLD.xls" 到 MATLAB.
[num, txt] =xlsread ('GLD');
% 第一列（从第二行开始）是交易日，格式为 mm/dd/yyyy
tday1=txt (2: end, 1);
% 将时间格式转化为 yyyymmdd
tday1=...
datestr (datenum (tday1,'mm/dd/yyyy'),'yyyymmdd');
% 将数据字符串转化为单元型变量，再转化为数值型变量
tday1=str2double (cellstr (tday1));
% 最后一列是调整后的收盘价
adjcls1=num (:, end);
% 读入 " GDX.xls" 到 MATLAB.
[num2, txt2] =xlsread ('GDX');
% 第一列（从第二行开始）是交易日，格式为 mm/dd/yyyy
tday2=txt2 (2: end, 1);
% 将时间格式转化为 yyyymmdd
tday2=...
datestr (datenum (tday2,'mm/dd/yyyy'),'yyyymmdd');
% 将数据字符串转化为单元型变量，再转化为数值型变量
tday2=str2double (cellstr (tday2));
adjcls2=num2 (:, end);
% 找出 GLD 或 GDX 有数据的所有交易日
tday=union (tday1, tday2);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
[foo idx idx1] =intersect (tday, tday1);
% 将两个价格序列放到一起
adjcls=NaN (length (tday), 2);
adjcls (idx, 1) =adjcls1 (idx1);
[foo idx idx2] =intersect (tday, tday2);
adjcls (idx, 2) =adjcls2 (idx2);
% 有数据缺失的交易日
baddata=find (any (~isfinite (adjcls), 2));
tday (baddata) = [];
adjcls (baddata,:) = [];
vnames=strvcat ('GLD','GDX');
% 用 ADF 检验执行协整检验
res=cadf (adjcls (:, 1), adjcls (:, 2), 0, 1);
prt (res, vnames);
% cadf 函数的输出结果:
% 两个变量的 ADF 检验:
% GLD, GDX
% CADF t-statistic        # of lags AR (1)        estimate
%        -3.35698533              1               -0.060892
%
%        1% Crit Value            5% Crit Value    10% Crit Value
%        -3.819                   -3.343           -3.042
% t-统计量为-3.36, 在 1%临界值-3.819 和 5%临界值-3.343 之间,
% 意味着有两个时间序列是协整的概率超过 95%
results=ols (adjcls (:, 1), adjcls (:, 2));
hedgeRatio=results. beta
z=results. resid;
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 对冲比率为 1.6766
% 即 GLD=1.6766 * GDX + z,
% 其中 z 是截距项，作为差价 GLD-1.6766 * GDX，应该是平稳的
% 画出的图同图 7-4 类似
plot (z);
```

如果你认为同行业中的任意两只股票都是协整的，我可以给出一个反例：KO（可口可乐）与 PEP（百事可乐）。用例 7.1 中一样的检验方法，我们可以发现，两只股票协整的可能性小于 90%。（你可以自己试试，并将结果与 epchan.com/book/example7_3.m 里的结果进行比较。）如果使用线性回归来拟合 KO 与 PEP，得到的时间序列图与图 7-5 相似。

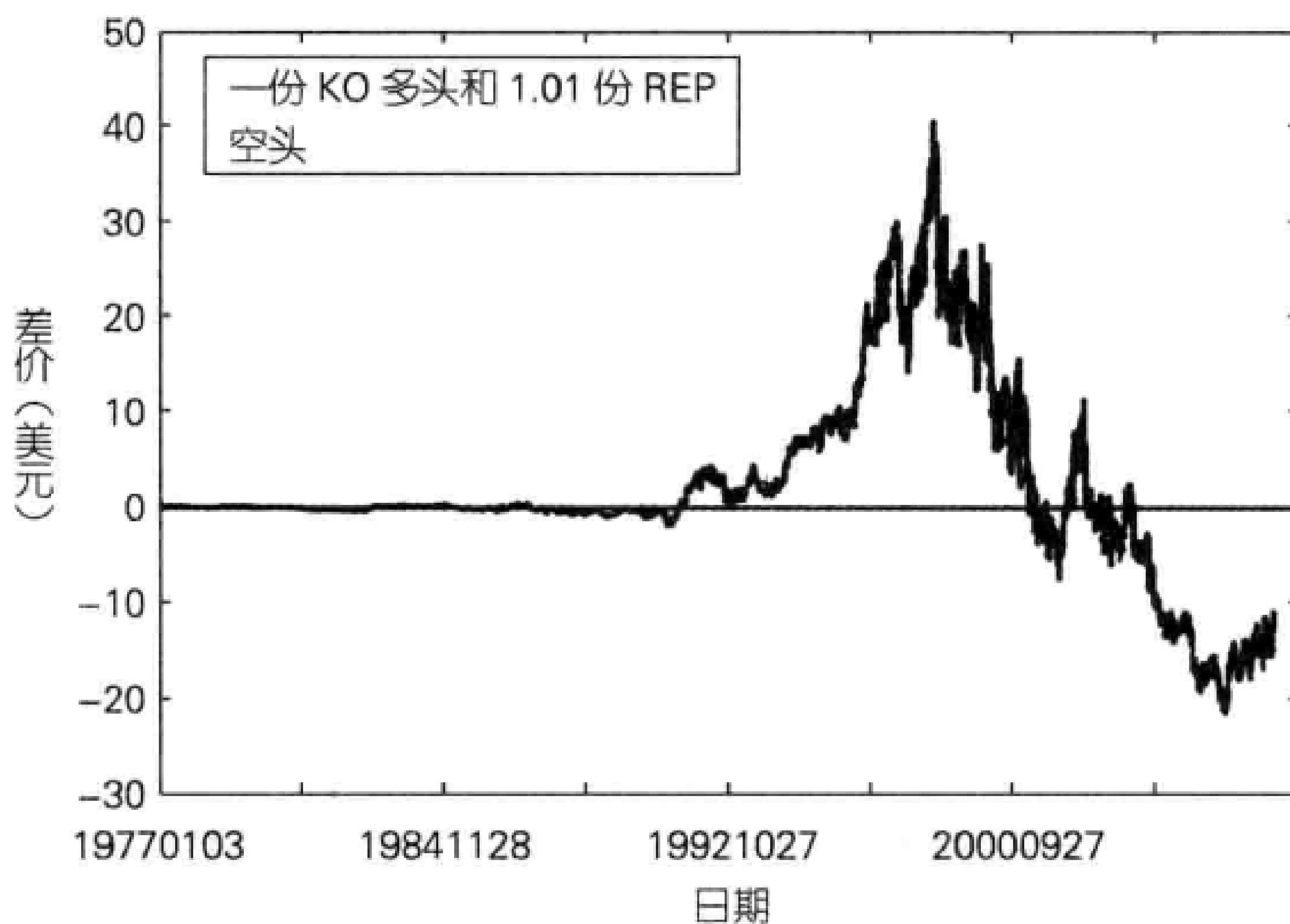


图 7-5　GLD 和 GDX 差价形成非平稳时间序列

如果一个价格序列（可以是一只股票、一对股票或是一个投资组合）是平稳的，只要未来继续保持平稳（这未必能保证），采用均值回归策略一定能盈利。反之则不然。成功的均值回归策略，并不一定要求一个平稳的价格序列。正如许多交易者所知道的，即使是一个非平稳的价格序列，

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

也可能有很多可以利用的短期回归机会。

许多配对交易员并不熟悉平稳性和协整性的概念。但他们中的大多数人都熟悉“相关性”，并且觉得与协整性似乎是同一个意思。实际上，这两个概念截然不同。两个价格序列的相关性实际上是指在一段时间（比如说 1 天）内其收益率的相关性。如果两只股票正相关，其价格变化方向在大多数时间都是相同的。然而，有正相关关系并不反映两只股票的长期行为特征。特别地，正相关无法保证两只股票在长期内价格偏离不会越来越大，即使在大多数时间里它们的变动方向相同。但是，如果两只股票在目前与未来都是协整的，它们的价格（经过合适的权重）就不大可能会偏离，即使它们的日（周或其他时间段）收益率可能不相关。

图 7—6 给出了两只虚构的股票 A 和 B，它们之间是协整的，但不相关。显然，股票 B 的价格变动与股票 A 明显不相关：有时它们的变动方向相同，有时相反。多数时间里，股票 B 的价格甚至根本没有变化。但是它们之间的差价总是会在一段时间后回到 1 美元左右。

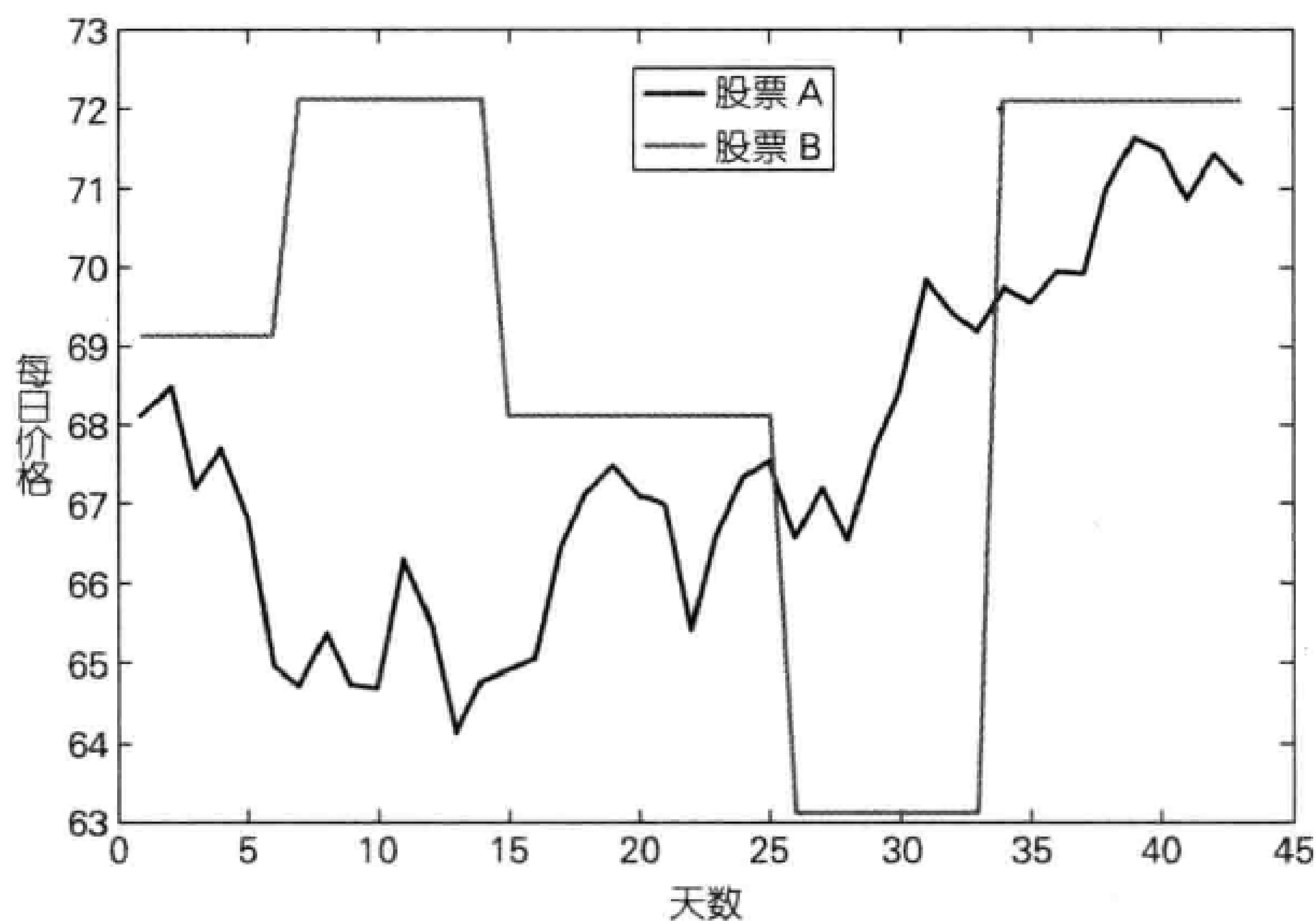


图 7—6　股票 A 和股票 B 是协整的但不相关

在现实中能找到这种现象的例子吗？当然，KO 与 PEP 就是。通过

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

程序"example7 3. m"可以知道，这两只股票不是协整的。然而，如果你对它们的日收益率的相关性进行检验，会发现它们之间的相关系数只有0.4849，从统计学上说显然是不相关的。例7.3给出了进行相关性检验的程序。

例7.3　检验KO和PEP之间的协整性和相关性

KO和PEP之间的协整性检验与例7.2中GDX与GLD之间的协整性检验类似，所以不再重复（可从epchan.com/book/example7_3.m下载）。协整性结果显示，ADF检验的t统计量为−2.14，大于10%的临界值−3.038，意味着这两个时间序列之间的协整概率小于90%。

下面给出了检验这两个时间序列相关性的代码：

```
% 相关性测试
dailyReturns= (adjcls−lag1 (adjcls)) ./lag1 (adjcls);
[R, P] =corrcoef (dailyReturns (2: end,:));
% R=
%
%      1.0000      0.4849
%      0.4849      1.0000
%
%
% P=
%
%      1      0
%      0      1
% P值为零意味着两个时间序列显著相关
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

平稳性并不仅限于股票之间的差价，汇率之间也可能存在平稳性。例如，加元/澳元（CAD/AUD）的交叉汇率就非常平稳。许多期货之间以及期货与固定收益证券之间都可能协整。（最简单的例子就是跨期差价组合：同一标的的商品、不同到期月份期货合约的多头和空头。固定收益证券业类似，同一发行人、不同期限债券的多头和空头。）

7.4 因子模型

财经评论员经常说“当前市场看好价值型股票”，“市场正在关注盈利增长”，或“投资者正关注通货膨胀数据”。我们应当如何量化这些及其他影响收益率的常见因素呢？

量化金融里有一个非常有名的因子模型（也称套利定价理论，APT），试图捕获影响收益率的各种因素，如盈利增长率、利率、市值等，这些因素被称为“因子”。数学上，可以将 N 只股票的超额收益率（收益率减去无风险利率）R 表示为：

$$R=Xb+u$$

其中，X 为 $N\times N$ 的因子风险矩阵，b 为 N 维因子收益率向量，u 为 N 维特有收益率向量。（虽然这一量化模型中的所有变量都与时间有关，但为了简便，这里我们假设都与时间无关。）

“因子风险”、“因子收益率”和“特有收益率”这些术语在量化金融模型中很常见，因此有必要先对这些概念做一些介绍。因子收益率是股票收益率的共同驱动因素，与单个股票无关。因子风险表示对各个共同驱动因素的敏感度。所有不能用共同因子解释的收益率部分就是特有收益率（比如，仅与某只股票有关，可以看做是 APT 模型中的随机噪声部分）。假设每只股票的特有收益率与其他股票的特有收益率之间是不相关的。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

我们通过 Fama-French 三因子模型（Fama and French，1992）这样一个简单的因子模型来说明。这一模型假定股票的超额收益率只与三个因子风险线性相关：股票的贝塔值（相对市场指数的敏感度）、市值、账面市值比率。对于不同股票或是不同时间段，这三个因子风险都是不同的。(常常将因子风险标准化，使得所有股票的因子风险的均值为 0，标准差为 1。)

现在，我们已经知道如何计算因子风险，那么因子收益率和特有收益率呢？我们无法直接计算因子收益率和特有收益率，必须用股票的超额收益率对因子风险进行多元线性回归，来得到因子收益率的估计值。注意，每只股票代表线性回归上的一个数据点，我们必须对每一个时间段单独做线性回归，如果需要多个时间段上的均值，可以将所有这些时间段上的值汇总成一个训练集，再对其做一次回归。

如果对 Fama-French 三因子模型进行线性回归来拟合多个时间段的数据，你会发现市值因子收益率通常是负的（意味着小盘股的收益率通常高于大盘股），账面市值比率因子收益率通常是正的（意味着价值型股票的收益率通常高于成长型股票）。由于大部分股票的业绩与市场指数是正相关的，贝塔因子收益率也是正的。

Fama-French 模型并非因子模型的唯一选择。事实上，你可以在模型中引入更多因子。例如，可以选择权益收益率作为一个风险因子，还可以选择股票收益率与基准利率的相关性作为另一个因子风险。你可以选择任意数量的经济的、基本面的或技术的因子。选择的因子风险是否合理将决定因子模型能否充分解释股票的超额收益率。如果因子风险（或者说整个模型）选择得不恰当，线性回归将会产生显著的特有收益率，并且拟合方程的 R^2 统计量将会很小。根据专家（Grinold and Kahn，1999）的建议，1 000 只股票的月度收益率和 50 个因子之间一个好的因子模型，R^2 应当在 30%～40%之间。

这些因子模型看起来好像只能解释历史数据，也就是说给定历史收益

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

率和因子风险，可以计算出历史时间段的因子收益率。但是，这些历史的因子收益率对我们的交易有何用处呢？事实上，因子收益率通常比个股收益率更稳定。换句话说，它们具有“惯性”。因此，可以假设它们的值从当期（从回归得到）到下一期保持不变。如果真是这样，只要因子风险选择得当并且时变的特有收益率不显著，就可以预测超额收益率。

这里需要做一点澄清。尽管我之前说，如果因子收益率具有惯性，因子模型就能够用于预测（因此也就能够用于交易），但并不意味着因子模型不能用于股票收益率的均值回归。实际上，因子风险可以用于均值回归。如果股票收益率是均值回归的，前一期的收益率为负值，则相应的因子收益率就会是正值。

如果你想建立一个基于基本面因子的交易模型，这里可以提供一些能够获取历史因子数据的供应商：

- Capital IQ：www. capitaliq. com
- Compustat：www. compustat. com
- MSCI Barra：www. mscibarra. com
- Northfield Information Services：www. northinfo. com
- Quantitative Services Group：www. qsg. com

例 7.4　因子模型举例：主成分分析

我之前提到的因子风险主要包括经济的（如利率）、基本面的（如账面市值比率）和技术的（如前期收益率）。获得一个包含大量股票的投资组合因子风险的历史数据，并用于对因子模型进行回测，对于独立交易员来说是非常昂贵且不切实际的。（那些准备购买数据的人，可以参考正文中的名单。）不过，有一种因子模型，其构建只依赖于历史收益率。这个方法叫做主成分分析（PCA）。

用主成分分析构建因子风险和因子收益率，必须假设因子风险在估计的时间段内是不变的（时间独立）。（这排除了表示均值回归或惯性的因

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

子，因为这些因子风险都与前期收益率有关）。更重要的是，如果假设因子收益率之间“不相关”，协方差矩阵 bb^T 就是对角矩阵。如果用协方差矩阵 RR^T 的特征向量作为 APT 方程 $R=Xb+u$ 中矩阵 X 的列向量，可知 bb^T 的确是对角矩阵，并且矩阵 RR^T 的特征值正好就是因子收益率 b 的方差。但是，如果因子数量与股票数量相等，我们就不需要使用因子分析了，因为只要选取几个具有较大特征值的特征向量就能构成矩阵 X。特征向量的个数是一个需要优化的交易模型参数。

下面的 MATLAB 程序（可从 epchan. com/book/example7_4. m 下载）展示了一个对 S&P600 小盘股使用主成分分析的可能交易策略。这一策略假设因子收益率具有惯性，即从本期到下期，因子收益率的值保持不变。因此，可以买入基于这些因子的期望收益率最高的股票，卖出期望收益率最低的股票。如果发现这一策略的平均收益率为负，表明对收益率具有惯性的假设是不合适的，或者策略的特有收益率太大了以至于策略失效。

```
clear;
% 使用回望交易日作为估计区间(训练集),以此来决定因子风险
% 回望期交易日为 252 天,因子 5 个
% 交易策略为:购买下一个交易日期望收益率最高的 50 只股票
topN=50;
% 选用 SP600 小盘股做测试(此 MATLAB 二进制输入文件包含交易日,
股票,开盘价,最高价,最低价,收盘价)
load('IJR 20080114');
mycls=fillMissingData(cl);
positionsTable=zeros(size(cl));
% dailyret 的行是在不同时间段上的观察值
dailyret=(mycls-lag1(mycls))/lag1(mycls);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
for t=lookback+1:length(tday)
% R 的列是不同的观测对象
R=dailyret(t-lookback+1:t,:)';
% 不考虑所有收益率缺失的股票
hasData=find(all(isfinite(R),2));
R=R(hasData,:);
avgR=smartmean(R,2);
% 移去均值
R=R-repmat(avgR,[1 size(R,2)]);
% 计算不同股票收益率的协方差矩阵
covR=smartcov(R');
% X 是因子风险矩阵,B 是因子收益率的方差
% 用 covR 的特征值作为 X 的列向量
[X,B]=eig(covR);
% 保留的因子数为 numFactors
X(:,1:size(X,2)-numFactors)=[];
% b 是从时间 t-1 到 t 的因子收益率
results=ols(R(:,end),X);b=results.beta;
% Rexp 是假设因子收益率保持常数时,下一个时间段的期望收益率
Rexp=avgR+X*b;
[foo idxSort]=sort(Rexp,'ascend');
% 做空期望收益率最低的 50 只股票
positionsTable(t,hasData(idxSort(1:topN)))=-1;
% 做多期望收益率最高的 50 只股票
positionsTable(t,...
hasData(idxSort(end-topN+1:end)))=1;
end
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 计算交易策略的每日收益率
ret=...
smartsum(backshift(1,positionsTable). * dailyret,2);
% 计算交易策略的年化收益率
avgret=smartmean(ret) * 252;% 收益率很低
% avgret =
%
% -1.8099
```

程序中使用了 smartcov 函数来计算多只股票日收益率向量的协方差矩阵。与 MATLAB 内置的 cov 函数不同，smartcov 函数忽略了收益率缺失的交易日（包括 NaN 值）。

```
function y = smartcov(x)
% n 个有限元素的协方差
% 行为观测值,列为变量
% 用 N 标准化,而非 N-1
y=NaN(size(x,2),size(x,2));
xc=NaN(size(x));
goodstk=find(~all(isnan(x),1));
xc(:,goodstk)=...
x(:,goodstk)-repmat(smartmean(x(:,goodstk),1),...
[size(x,1)1]);% 移去均值
for m=1:length(goodstk)
    for n=m:length(goodstk)
    y(goodstk(m),goodstk(n))=...
      smartmean(xc(:,goodstk(m)).
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
    *..xc(:,goodstk(n)));
    y(goodstk(n),goodstk(m))=y(goodstk(m),goodstk(n));
    end
end
```

因子模型在实际交易中的业绩如何呢？其实这主要取决于所使用的因子模型。但总体上看，基本面和宏观经济因子主导的因子模型有一个主要缺陷：假设投资者始终会用相同的标准对公司进行估值。换句话说，因子收益率必须具有惯性，模型才会有效。

例如，虽然价值（账面市值比率）因子的收益率通常是正的，但有时候会是负值，如 20 世纪 90 年代后期的互联网泡沫期间，2007 年 8 月和 12 月，由于投资者偏好成长性股票的原因。就像《经济学家》报道的那样，成长型股票在 2007 年受到追捧的一个原因就是相对于价值型股票的溢价已经大幅度缩小了（Economist，2007b）。另一个原因是美国经济增长的放缓，使得投资者倾向选择那些盈利能够持续增长，而不是那些受经济衰退所拖累的股票。

因此，在投资者的估值方法发生改变的这段时间内，因子模型也很可能出现大幅度挫跌，即使是很短的一段时间。不过话说回来，对于任何隔夜持仓的交易模型来讲，这是个常见的问题。

7.5 清仓策略

每个策略都有其特定的建仓信号，但是策略的清仓信号并无太大区别，主要有以下几种：

- 固定的持有期
- 目标价格或盈利上限

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

- 最新的建仓信号
- 止损价格

无论是惯性模型、回顾模型，还是基于惯性或回归的季节性交易策略，固定持有期都是其默认的清仓策略（后面将详述）。前面说过，信息的缓慢扩散是惯性产生的一个原因，这种情况下，这个过程的期限有限。通过回测可以获得这个期限的平均时间，这一时间决定了最优持有期的长度。

在确定惯性模型的最优持有期时需要注意：最优持有期会随信息扩散速度的加快以及使用同一策略的交易员人数的增加而缩短。因此，一个回测期间的最优持有期为一周的惯性模型，实际中可能只有一天持有期。更糟糕的是，一年后这一策略就可能无法盈利了。同时，由于历史数据的数量有限，回测时难免会存在数据迁就偏差。不幸的是，对于新闻或事件驱动的惯性策略，没有其他的清仓选择。对于均值回归策略，我们有更加具有统计可靠性的方法来决定交易的最优持有期。

时间序列的均值回归模型可以运用 Ornstein－Uhlenbeck 公式（Uhlenbeck，1930）来建模。用符号 z（t）来表示股票配对的均值回归差价（多头的市场价值减去空头的市场价值）。有：

$$dz(t) = -\theta(z(t) - \mu)dt + dW$$

其中，μ 为价格的平均价值，dW 为某种高斯随机噪声。给定日差价的时间序列，将差价对每日变化 dz 进行线性回归，很容易得到 θ（和 μ）。数学上可知，$z(t)$ 的均值服从均值为 μ 的指数衰减过程，其半衰期（差价回归到最初偏离均值的一半所需要的期望时间）为 $ln(2)/\theta$。半衰期可以用来确定均值回归头寸的最优持有期。由于所有的历史时间序列数据（而不仅仅是交易促发的那些天）都可以被用来估计 θ，通过这种方法估计的半衰期比直接从交易模型中获得的半衰期更加可靠。例 7.5 依然使用 GLD 和 GDX 价差来估计均值回归的半衰期。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例 7.5　计算均值回归时间序列的半衰期

我们可以通过例 7.2 中 GLD 和 GDX 的均值回归差价来计算均值回归半衰期。MATLAB 代码可以从 epchan.com/book/example7_5.m 获得。(这个程序的第一部分与 example7 2.m. 相同。)

```
% = = = 在此插入 example7_2.m = = =
% = = = Insert example7 2.m in the beginning here = = =
prevz=backshift(1,z);% z at a previous time-step
dz=z-prevz;
dz(1)=[];
prevz(1)=[];
% 假设 dz=theta*(z-mean(z))dt+w,
%w 是扰动项
results=ols(dz,prevz-mean(prevz));theta=results.beta;
halflife=-log(2)/theta
% halflife =
%
% 10.0037
```

GLD 和 GDX 的均值回归半衰期大约为 10 天，这就是在策略盈利之前应该持有该差价的期望天数。

如果你认为所持有的证券是均值回归的，就得计算一下你的目标价格——该证券历史价格的均值，也就是 Ornstein-Uhlenbeck 公式的 μ。这一目标价格可以与半衰期一同作为清仓信号使用（即只要满足一个条件就清仓）。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

如果我们拥有公司的基本面估值模型，按目标价格清仓的方法也能用于惯性模型。但是，基本面估值并非很科学，所以，目标价格在惯性模型中的使用并不像在均值回归模型中那样靠谱。如果使用基于基本面估值的目标价格可以轻松获利，所有投资者必须做的是，通过每日研读股票研究报告做出投资决策。

假定你正运行一个交易模型，并根据信号建仓。一段时间后，你再次运行模型。如果模型给出了相反的交易信号（比如：当你拥有一个空头头寸时，系统发出了“买入”的信号）。此时你有两个选择：要么清仓，要么清仓后反向建仓。无论哪种方法，实际上都是使用交易模型最新的建仓信号来作为原仓位的清仓信号。当一个交易模型只能在最优持有期中的一个更短区间内运行良好时，这种方法是获得清仓信号的常见方式。

注意，这种清仓策略同时也提供了一个止损策略建议。在惯性模型中，如果模型最近发出的建仓信号与现有头寸方向相反，惯性的方向很可能就已经改变，你的头寸已经发生了亏损（或者更准确地说是挫跌）。此时，清仓就相当于进行止损。然而，对于惯性模型来说，将最新的反向建仓信号作为止损信号，显然比按照可人为调整的参数而确定的任意止损价格来得更合理，因为引入参数的过程难免会存在数据迁就偏差。

当我们使用回归模型时，有必要考虑一个类似的状态。如果现有头寸处于亏损状态，回归模型给出的交易方向必然还是一样的。所以，回归模型的建仓信号永远不会建议止损。（相反地，它可以给出一个目标价格或者盈利上限，此时，回归力度会很大以至于发出了相反方向的建仓信号。）其实，在一个均值回归模型中，基于持有期或盈利上限的清仓方式比通过止损清仓更加合理，因为止损就意味着你在更加极端的情况下清仓，这与模型的假设相违背。（除非你相信，由于新消息的影响，市场已经由均值回归状态转变为惯性状态了。）

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.6 季节性交易策略

季节性交易策略也称作“日历效应”。一般而言，这种策略建议在每年的固定日期买入或卖出某种证券，并在另一个固定日期平仓。它既可应用于股票市场，也可应用于商品期货市场。也许是因为这种交易机会已广为所知，就我个人看来，目前很多股票市场上的季节性交易策略都已失效。不过，在商品期货市场上仍然还有部分季节性交易策略是可以盈利的。

股票市场最著名的季节性交易策略是“一月效应”。这一策略有许多种版本，其中一个版本是：前一年收益率最差的那些小盘股，在第二年一月的收益率会普遍高于前一年业绩最好的那些小盘股（Signal，2006）。可能是因为税收原因，投资者喜欢在 12 月清算亏损，这会导致价格下行压力。当这种压力在一月消失的时候，价格就会有一定程度的恢复。虽然在 2006—2007 年间这一策略的业绩不尽如人意，但是 2008 年的业绩相当完美。对于均值回归而言，2008 年 1 月是一个值得铭记的月份（在这个月，法国兴业银行的特大交易丑闻间接导致了美联储在开市前紧急降息 75 个基点。由此带来的市场动荡击垮了许多惯性交易者，但是均值回归策略却在最初的大额挫跌、之后又在受美联储影响戏剧般回升的过程中，获得了巨大收益）。例 7.6 给出了一月效应策略的回测代码。

例 7.6　回测一月效应

这里给出了在标准普尔 600 小盘股上运用一月效应策略计算收益率的 MATLAB 代码。（源代码和数据可以从 epchan.com/book/example7_6.m 下载。）

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
clear;
load('IJR 20080131');
onewaytcost=0.0005;% 5bp one way transaction cost
years=..
year(cellstr(datestr(datenum(cellstr(..
num2str(tday)),'yyyymmdd'))));
months=..
month(cellstr(datestr(datenum(cellstr(..
num2str(tday)),'yyyymmdd'))));
nextdayyear=fwdshift(1,years);
nextdaymonth=fwdshift(1,months);
lastdayofDec=find(months==12 & nextdaymonth==1);
lastdayofJan=find(months==1 & nextdaymonth==2);
% lastdayofDec 从 2004 年开始，
% 所以将 2004 从 lastdayofJan 中删掉
lastdayofJan(1)=[];% 确定每个 lastdayofDec 后的 lastdayofJan
assert(all(tday(lastdayofJan)> tday(lastdayofDec)));
eoy=find(years? =nextdayyear);% 年末指数
eoy(end)=[];% 上一个指数不是年末的
% 确保 eoy 日期与 lastdayofDec 日期匹配
assert(all(tday(eoy)==tday(lastdayofDec)));
annret=..
(cl(eoy(2:end),:)-cl(eoy(1:end-1),:))./..
cl(eoy(1:end-1),:);% 年收益率
janret=..
(cl(lastdayofJan(2:end),:)-
cl(lastdayofDec(2:end),:))./cl(lastdayofDec(2:end),:);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 1 月收益率
for y=1:size(annret,1)
% 选取年收益率可观的股票
hasData=..
find(isfinite(annret(y,:)));
% 根据前一年的收益率排序
[foo sortidx]=sort(annret(y,hasData),'ascend');
% 买入十分位数(decile)收益率最低的股票，卖出十分位数收益率最高的
股票
topN=round(length(hasData)/10);
% 组合收益率
portRet=..
(smartmean(janret(y,hasData(sortidx(1:topN))))-..
smartmean(janret(y,hasData(..
sortidx(end-topN+1:end)))))/2-2 * onewaytcost;
fprintf(1,'Last holding date %i:Portfolio
return=%7.4f\n',tday(lastdayofDec(y+1)),portRet);
end
% 这里应该是输出
% 最后持有期 20051230:组合收益率=-0.0244
% 最后持有期 20061229:组合收益率=-0.0068
% 最后持有期 20071231:组合收益率= 0.0881
```

这一程序用到了大量的工具函数。第一个是 assert 函数，用于确保程序是正确的。

```
function assert(pred,str)
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 若前提条件非真,ASSERT 会报错.
% assert(pred,string)
if nargin<2,str =';end

if~pred
s = sprintf('assertion violated:%s',str);
error(s);
end
```

第二个函数是 fwdshift 函数,与 lag1 函数相反,是将时间序列函数向前移动一步。

```
function y=fwdshift(day,x)
assert(day>=0);
y=[x(day+1:end,:,:);..
NaN * ones(day,size(x,2),size(x,3))];
```

最近，有人提出了另外一个股票季节性策略（Heston and Sadka, 2007；可从 lcb1. uoregon. edu/ rcg/ seminars/ seasonal072604. pdf 下载)。这一策略非常简单，就是每个月买入大量去年这个月业绩最好的股票，同时卖空同样数量的去年这个月业绩最差的股票。如果不计交易成本，这一策略在 2002 年之前的平均年收益率超过 13%。但是我发现这一策略在最近似乎失效了，你可以通过例 7. 7 对这个策略进行测试。(见我博客的读者评论：epchan. blogspot. com/ 2007/ 11/ seasonal – trades – in – stocks. html。)

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

例 7.7　回测跨年季节趋势策略

这里给出了我之前提到的跨年季节趋势交易策略的 MATLAB 代码。(源代码可从 epchan.com/book/example7_7.m 下载。) 由于数据是基于 2007 年 12 月 23 日获取的标准普尔 500 指数，所以存在存活偏差。

```
clear;
load('SPX 20071123','tday','stocks','cl');

%找到月底几天的指数.
monthEnds=find(isLastTradingDayOfMonth(tday));
monthlyRet=..
(cl(monthEnds,:)-lag1(cl(monthEnds,:)))./..
lag1(cl(monthEnds,:));
mycl=fillMissingData(cl);
% 按月收益率升序排列
[monthlyRetSorted sortIndex]=sort(monthlyRet,2);
% 排序后的前一年本月收益率
prevYearMonthlyRetSorted=backshift(12,monthlyRetSorted);%去年排
序后的指数
prevYearSortIndex=backshift(12,sortIndex);
positions=zeros(size(monthlyRet));
for m=13:size(monthlyRet,1)
hasReturns=...
isfinite(prevYearMonthlyRetSorted(m,:))&..
isfinite(cl(monthEnds(m-1),:));
mySortIndex=prevYearSortIndex(m,hasReturns);
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 买入十分位数最高的几只股票,卖出十分位数最低的几只股票
topN=floor(length(mySortIndex)/10);
positions(m-1,mySortIndex(1:topN))=-1;
positions(m-1,..
mySortIndex(end-topN+1:end))=1;
end
ret=smartsum(lag1(positions).*monthlyRet,2);
avgannret=12*smartmean(ret);
sharpe=sqrt(12)*smartmean(ret)/smartstd(ret);

fprintf(1,..
'Avg ann return=%7.4f Sharpe ratio=%7.4f\n',..
avgannret,sharpe);
% 输出应该是年收益率均值=-0.9167 Sharpe ratio=-0.1055
```

这一程序用到了几个工具函数,第一个是 LastTradingDayOfMonth。它返回 1 和 0 的逻辑数组,用以判断某个交易日是否是一个月的最后一个交易日。

```
function isLastTradingDayOfMonth=..
isLastTradingDayOfMonth(tday)
% isLastTradingDayOfMonth=isLastTradingDayOfMonth(tday)返回一个逻辑序列。如果 tday(t)确为当月的最后交易日,返回 True。
tdayStr=datestr(datenum(num2str(tday),'yyyymmdd'));
todayMonth=month(tdayStr);
tmrMonth=fwdshift(1,todayMonth);% tomorrow's month
isLastTradingDayOfMonth=false(size(tday));
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
isLastTradingDayOfMonth(todayMonth? =tmrMonth & ..
isfinite(todayMonth)& isfinite(tmrMonth))=true;
```

另一个工具函数是 backshift 函数，它与 lag1 函数类似，但可以平移任意天数。

```
function y=backshift(day,x)
% y=backshift(day,x)
assert(day>=0);
y=[NaN(day,size(x,2),size(x,3));x(1:end-day,:,:)];
```

如果你尝试最近 5 年而不是全部数据，会发现平均收益率更低。

相比股票的季节性交易策略，商品期货市场的季节性交易策略仍然是有利可图的。这也许是因为对某种商品的需求来源于实体经济而非投机。

汽油期货交易是最直观的商品季节性交易策略之一：只需要在临近 4 月中旬买入 5 月份到期的汽油期货合约，并在 4 月下旬卖出就可以了。直到 2008 年 4 月写作本书之时，这一交易在之前的 11 年里都是盈利的（详情见“汽油期货的季节性交易”）。似乎北美的汽油期货价格在临近夏季自驾高峰前总会有一段上扬。

汽油期货的季节性交易

在夏季自驾游高峰到来时，汽油期货价格上涨并没有什么奇怪的。对于一名交易员而言，需要关心的是应该买入哪个月份的合约，并持有多长时间。通过阅读许多文献，我发现到目前为止，主要的策略是在 4 月 13 日（如遇节假日，则顺延至下一个交易日）收盘时买入 1 份 RB（纽约商品交易所的无铅汽油期货合约），并在 4 月 25 日收盘时卖出（如遇节假

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

日，则提前至前一个交易日）。从历史业绩来看，这一策略自 1995 年以来都是盈利的，下表给出了这一交易头寸（2007—2008 年的数据来自于实际交易）每年的盈亏和最大挫跌（从建仓的第 1 天开始计算）：

年份	盈亏（美元）	最大挫跌（美元）
1995	1 037	0
1996	1 638	−2 226
1997	227	−664
1998	118	0
1999	197	−588
2000	735	−315
2001	1 562	0
2002	315	0
2003	1 449	−38
2004	361	−907
2005	6 985	−25
2006	890	0
2007*	2 286	−9 816
2008*	4 741	0

*2 倍 QU 的实际交易结果

对于那些不愿意冒太大风险的交易员，可以选择买入纽约商品交易所的微型汽油期货 QU，这种合约的大小只是 RB 合约的一半，尽管其流动性要差一些。

（本研究的灵感来源于 Paul Kavanaugg 在 PFBest.com 上发表的月度季节性交易文章。你可以在 Fielden（2005）或 Toepke（2004）上读到这一文章以及其他季节性期货交易模式的内容。）

除了汽油外，出于空调用电的需要，发电厂在临近夏季时对天然气的需求也会增加。因此，另一个曾连续 13 年获利的商品季节性交易策略与

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

天然气有关：在临近 2 月底时买入 6 月到期的天然气合约，并在 4 月中旬卖出。(详情参见下表。)

天然气期货的季节性交易*

出于空调用电的需要，发电厂在夏季对天然气的需求会增加。这就意味着我们可以建立这样一个天然气季节性交易策略：在 2 月 25 日（如遇节假日，则顺延至下一个交易日）收盘时买入一份 6 月到期的纽约商品期货交易所天然气合约（代码：NG），并在 4 月 15 日（如遇节假日，则提前至前一个交易日）清仓。在写作本书时，这一策略已经连续 14 年盈利。下表给出了其每年的盈亏和最大挫跌（回测和实际交易）：

年份	盈亏（美元）	最大挫跌（美元）
1995	1 970	0
1996	3 090	−630
1997	450	−430
1998	2 150	−1 420
1999	4 340	−370
2000	4 360	0
2001	2 730	−1 650
2002	9 860	0
2003	2 000	−5 550
2004	5 430	0
2005	2 380	−230
2006	2 250	−1 750
2007	800	−7 470
2008*	10 137	−1 640

* 2 倍 QG 的实际交易结果

天然气合约的波动巨大，曾导致对冲基金（如 Amaranth Advisors 亏损 60 亿美元）和银行巨头（如 Bank of Montreal 亏损 4.5 亿美元）的巨

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

额亏损。所以，一定要谨慎使用这一策略，也许可以考虑使用微型 QG 期货合约，这个合约的大小只是 NG 合约的一半。

*这篇文章最早发表于我的网站 epchan.com/subscription，这里更新了数据。网站的用户名和密码为“sharperatio”。

虽然连年盈利，商品期货季节性交易策略也有其缺陷：通常一年只能做一次，所以我们无法判断回测结果是否是由数据迁就偏差引起的。通常，减少这一问题的方法之一是，通过选择不同的建仓和清仓时间来考察策略的盈利能力是否稳定。此外，交易者必须记住，只能交易那些有实际经济意义的季节性交易策略，如上面提到的汽油和天然气的季节性交易。

7.7 高频交易策略

一般而言，如果策略的目标是获取高夏普比率（其实也应该是这个，正如第 6 章所讨论的），交易应当采取高频而非隔夜持仓。

什么是高频交易策略？为何它会有更高的夏普比率？许多高频交易专家认为高频交易策略是那些持仓不超过几秒的策略，而本书采用的是更广泛的定义，即把日内交易策略也归为高频交易策略。由于充足的流动性，高频交易策略最早出现在外汇市场，随后是期货市场。最近六七年，随着股票市场流动性增强，逐笔历史数据普及以及计算速度大幅提高，高频交易策略也已被广泛应用于股票交易。

高频交易策略能获得高夏普比率的理由很简单：根据大数定律，交易的次数越多，收益率相对于均值的偏差就越小。而在高频交易策略下，一天可交易成百上千次。因此，如果你的策略具有正的平均收益率，高频交易日收益率与平均收益率的偏差将会达到最小。由于高频交易策略具有更高的夏普比率，与非高频策略相比，可使用的杠杆水平更高，高杠杆进而

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

大大提高策略的净值收益率。

当然，大数定律并没有解释为什么某个高频交易策略会有正的平均收益率。实际上，重要的是我们只需要解释为什么高频策略在总体上是盈利的就行了，因为目前这个策略已经被市场上很多基金经理所采用。有均值回归策略，也有惯性策略；有市场中性的交易员，也有多头单向的交易员。但通常来说，这些策略要么是利用市场中出现的微小的无效性而获利，要么是通过提供短暂的流动性以获取微小的报酬。与基于宏观趋势或公司基本面的策略会时常经历大的波动不同，高频交易策略依赖于每天都会出现的市场短暂无效和流动性需求，这使得它可以每天持续盈利。此外，高频交易策略要求交易量适中，因此进行风险管理非常容易：亏损时可以很快地“去杠杆”，市场糟糕时可随时停止交易。最坏的结果也不过是当某个策略太过普通，收益率不断下滑而最终无利可图。至于突发大额亏损或多账户的连环亏损则是不太可能的。

虽然成功的高频交易策略有许多优点，但是当平均持有期只有几分钟甚至几秒钟的时候，对其进行回测就不那么容易了。对这种策略进行回测，交易成本对最终结果会产生重大影响。在不考虑交易成本的情况下，即使是最简单的高频交易策略也是能够盈利的。因此，仅仅有最新的高频数据远远不够，还需要考虑买价、卖价以及最新的报价，才能比较按买价成交与按卖价成交的盈利水平。有时，回测还需要历史指令下单信息。除非具有十分高级的模拟器，否则最好的检测方法还是实时交易。

当然，回测只是高频交易的很小一部分，而高速执行往往决定交易的实际盈亏。专业的高频交易公司都是用 C 语言而不是其他用户友好性更高的编程语言来编写策略，并且会将服务器放置于交易所或主要网络节点附近以降低时滞。所以，虽然高频交易策略夏普比率较高并且收益率可观，但还是不大适合初涉此领域的独立交易员。当然，随着交易经验和资源的积累，我们没有理由不向这个目标前进。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

7.8 高杠杆组合优于高贝塔组合吗?

第 6 章讨论了基于凯利公式的应用于投资组合的最优杠杆水平。在有关因子模型的那一节，讨论了 Fama－French 三因子模型，从中我们知道投资组合（或股票）的收益率与它的贝塔值（如果股票的市值和账面价值是固定的）成正比。换句话说，我们可以通过提高杠杆水平或提高贝塔值（选择高贝塔值股票）来提高投资组合的收益率。这两种方法似乎都合乎常理。事实上，可以通过提高一个低贝塔值投资组合的杠杆使其收益率与高贝塔值投资组合的收益率相当。假设两个投资组合具有相同的平均市值和账面价值，根据 Fama－French 模型，这两个投资组合应该有相同的平均收益率（忽略特有收益率，可以通过增加投资组合中的股票数量来降低）。问题是，高杠杆低贝塔的投资组合与低杠杆高贝塔的投资组合是等价的吗?

回答是否定的。第 6 章讨论过，在使用凯利杠杆的条件下，投资组合的长期复合增长率与夏普比率的平方成正比，而不是与平均收益率成正比。所以，如果两个投资组合的平均收益率相同，我们应当选择风险或标准差较小的那个组合。而实证研究表明，由低贝塔值股票构成的投资组合往往风险较低、夏普比率较高。

例如，PanAgora 资产管理公司的 Edward Qian 博士在《风险对等组合》(未公开发表）中写道，股票和 60－40 的资产配置并不是最优的，因为它的风险资产（股票）比重过大。如果想在保持风险水平不变的情况下，实现更高的夏普比率，Qian 博士建议 23－77 的股票和债券配置，并使用 1.8 倍的杠杆。

不知何故，市场经常低估高贝塔值股票。因此，对高贝塔值股票的投资组合和低贝塔值股票的投资组合进行选择时，应选择低贝塔值股票的投

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

资组合，并通过提高杠杆来获得最大的复合增长率。

不过要注意，所有这些结论都是基于收益率服从正态分布的假设（见第 6 章的相关讨论）。因为真实收益率的分布具有厚尾特征，对低贝塔值股票使用过高的杠杆并不是明智的。

7.9 小 结

本书主要介绍量化交易的一种特定类型："统计套利"。虽然名字很专业，但是无论从概念上看还是从数学计算上看，统计套利实际上还是要比衍生品（如期权）交易或固定收益工具交易简单。我已经介绍了统计套利的大部分标准做法：均值回归和惯性、状态转换、平稳性与协整性、套利定价理论或因子模型、季节性交易模型、高频交易。

现将一些应注意的重点总结如下：

- 均值回归状态比趋势状态更常见。
- 回测均值回归时存在一些棘手的数据问题，如异常报价和存活偏差。
- 新信息的缓慢扩散、机构投资者大额指令的成交或"羊群效应"会引发趋势状态。
- 交易员之间的竞争会减少均值回归交易的机会。
- 交易员之间的竞争会缩短惯性交易的最优持有期。
- 用数据挖掘方法可以发现状态转换。
- 平稳价格序列对均值回归交易非常理想。
- 两个或更多的非平稳性价格序列只有是协整的时候才可以构成一个平稳序列。
- 协整性与相关性不同：协整性考察的是两只或多只股票"价格"的长期关系，而相关性与短期收益率有关。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

- 因子模型，或套利定价理论，常用于模拟基本面因素对股票收益率的线性影响。
- 最著名的因子模型之一是 Fama－French 三因子模型，它假设股票收益率与贝塔值和账面市值比率正相关，与市值负相关。
- 因子模型会因为状态转换而有一个相对较长的持有期和挫跌期。
- 均值回归策略和惯性策略清仓信号不同。
- 均值回归策略最优持有期估计由于 Ornstein－Uhlenbeck 公式而比较可靠。
- 惯性策略最优持有期估计容易因为信号数量较少而出错。
- 止损适用于惯性策略，但不适用于均值回归策略。
- 近年来，股票的季节性交易策略（如日历效应）已经变得无利可图。
- 商品期货的季节性交易策略仍然可以继续盈利。
- 根据“大数定律”，高频交易策略具有较高的夏普比率。
- 较高的夏普比率使高频交易策略具有最高的长期复合增长率。
- 对高频交易策略进行回测非常困难，策略执行有技术依赖。
- 高杠杆的低贝塔值股票组合的长期复合增长率，比低杠杆的高贝塔值股票组合高。

大多数统计套利交易策略都是这些效应或模型的组合。这些策略能否盈利，更多的与使用它们的地点和时间有关，与策略本身在理论上是否正确无关。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

第 8 章　结语：独立交易员能否成功?

量化交易在 2007 年的夏天变得臭名昭著。一些由最负盛名的投资经理掌控的大型对冲基金在短短数日内亏损了几十亿美元（虽然一些基金在月底挽回了亏损）。这使人们不愉快地回想起了其他声名狼藉的对冲基金，如长期资本管理公司和 Amaranth Advisors（第 6 章都有所提及），只是这次不是一名交易员或一家公司发生问题，而是许多基金在短期内同时遭受重创。

早在我从事机构量化交易职业开始，就与许多独立交易员交流过。我发现，他们通常在简陋的办公室或卧室工作，并且全年赚取少量但稳定增长的盈利。这可能与公众设想的交易员的典型形象相去甚远。事实上，我认识的许多独立交易员不但在大基金遭受巨额亏损时存活了下来，甚至在那段时间里还盈利颇丰。多年以来，我一直对这一现象迷惑不解：为什么这些独立交易员净值微不足道、设备简陋，却能够获得高夏普比率，而那些拥有明星团队的公司反而失败了?

2006 年初，我从机构资产管理部门离职，开始了自己的独立交易员经历。我想知道自己在没有机构约束和规定的环境下交易是否还能盈利，同时也想弄清楚独立交易究竟是不是一场骗局，以及我到底适不适合做一

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

名独立交易员。那时我向自己保证，只要遇到上述的任何问题，这辈子就永远不再交易了。幸运的是，我最终还是存活了下来，并且在这一过程中，还找到了之前那个一直令我疑惑的答案。

事实上，关键就是“容量”，这个概念在第 2 章已经介绍过了（重复一下：容量就是一个策略能产生较好收益率的净值规模）。相比于 1 亿美元的账户，10 万美元的账户获得较高夏普比率要容易得多。许多简单且可盈利的策略对小容量有效，但对于大型对冲基金，这些策略可能都不适用。这就给我们独立交易员提供了生存空间。

有必要进一步讨论一下容量的问题。许多盈利的小容量策略其实就是在扮演做市商角色：在市场需要短期流动性时提供流动性，在流动性需求消失时获利走人。但是，如果管理着几十亿美元，你变成了需要流动性的一方，就不得不为获取流动性而支付一定的成本。为最小化流动性需求的成本，你只能持有头寸较长时间。而持有头寸较长时间，你的投资组合就会遭遇不利的宏观变化（如状态转换）。尽管模型合理的话，长期可以盈利，但是那些偶然事件导致的大幅挫跌仍无法避免，而这样的事往往会成为头条新闻。

大型机构的大容量策略还有其他劣势。对冲基金间的激烈竞争意味着策略的几无盈利。过低的收益率迫使基金经理不得不过度使用杠杆。为了在竞争中获胜，交易员们必须开发更加复杂的模型，这又容易产生数据迁就偏差。尽管交易模型越来越复杂，但他们努力寻找的市场无效性仍可能是一样的，从而最终还是持有非常相似的头寸。我们在第 6 章讨论过这种现象。当市场环境改变时，相似损失头寸争先恐后的清仓，可能会（实际上确实会）引起市场的彻底崩盘。

独立交易员在大型基金失败的时候能够成功的另一个原因是，基金的管理层施加的各种各样约束。例如，量化基金的交易员可能被禁止使用纯多头策略，而纯多头策略往往更容易发现、更加简单、盈利更高，如果小规模交易，也不会比市场中性策略的风险更大。或可能被禁止从事期货交

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

易，或可能被要求不仅要市场中性还要行业中性。当你认为均值回归策略有效时，可能被要求去寻找一个惯性策略。诸如此类。不可否认，有些约束是出于风险管理的需要，但有些可能是管理层的一时心血来潮，或莫名其妙的想法。所有学过数学最优化的学生都知道，任何施加在最优化问题上的约束条件都会降低最优值。同样，基金施加的约束也同样会降低策略的收益率，如果没有同时影响到夏普比率的话。最后，有些监督量化基金前台投资经理的高管，由于并不精通量化技术，他们往往不是根据量化理论做出决策。

当策略表现出盈利迹象时，这些管理层就会施加强大压力，要求你迅速扩大投资规模。一旦策略开始出现亏损，他们又可能会要求你马上清仓，弃用这一策略。这些量化投资中的干预行为没有一个是数学上的最优。

除此之外，这些管理层人员的性格往往反复无常，这又对量化投资造成不利影响。发生亏损时，他们最缺的往往就是理智。

而独立交易员可以免受这些限制和干预。只要你性格坚定，能够坚持量化交易的基本原则，你的交易环境可以比大型基金更加接近最优。

事实上，还有一个原因可以解释为什么对冲基金比独立交易员更容易失败。当一个人用其他人的资金进行交易时，好的结果几乎是无限制的，而最坏的结果也就是被炒鱿鱼而已。这就导致在遵守机构严格的常规风险管理约束之外，只要能够避开风险管理者的监督，交易员就基本上都会趋向于采用高风险的策略。法兴银行的热罗姆·凯维埃尔先生就向我们证实了这一点。

这一事件最终给法兴银行造成了71亿美元的亏损，并间接导致了美联储的紧急降息。银行内控机制在三年中都没有发现这一欺诈交易的原因，在于凯维埃尔先生曾在银行后台工作过，非常清楚如何避开内控机制(Clark，2008)。

其实凯维埃尔先生的欺诈技巧并非原创。我曾供职于一家大型投资银

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

行，那里有两位进行量化交易的自营交易员。他们喜欢躲在交易大厅某个角落的玻璃窗下进行交易。一方面是他们觉得这样就可以不受那些聒噪的非量化交易员打扰，另一方面是认为这样可以保守他们的交易秘密，好吧，秘密。据我所知，他们几乎不与任何人交谈，甚至他俩之间也似乎从未有过交流。

有一天，其中一名交易员消失了，并且再也没回来。很快，大批审计人员检查了这名交易员的文件和电脑。结果发现，就像凯维埃尔先生一样，这个曾在 IT 部门工作过、擅长计算机操作的交易员成功伪造了上百万美元的虚假盈利，却从未被任何人发现，直到有一天，某台电脑的崩溃导致他的欺诈程序停止运行，从而暴露了他的行为。有传言说他早已逃到了印度，在那里享受着高品质的生活。

你瞧，就是这么回事。我希望通过这些可信的案例让你相信，只要在交易过程中严守纪律并用心去做，独立交易员比机构交易员更有优势。当然，作为独立交易员的附带好处是，从一开始就是自由的。就个人而言，我觉得目前的工作状态要比之前开心得多，当然这并不包括那些难以避免且惊心动魄的遭遇挫跌的日子。

8.1 接下来

假设你已经找到了一些既好又简单的策略，并正在家里开心地交易。接下来要做什么？如何实现“增长”？

其实，第 6 章已经讨论过增长，不过是狭义上的。用凯利公式能够实现净值的指数增长，这一增长的上限是策略的容量。之后，增长来自于策略数量的增加。例如，你可以寻找比现有策略交易频率更高的策略。这样做的话，投资是必不可少的，升级设备，购买更昂贵的高频历史数据。或者与之相反，你也可以去寻找那些有更长持有期的策略。尽管夏普比率较

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

低，但可以显著提高策略的容量。不过这些策略大多需要使用价格昂贵的历史基本面数据进行回测。如果你是一名股票交易员，可以尝试进行期货或外汇交易，这些领域的容量通常比股票高。如果你想不出好点子或对于想要进入的新市场缺乏经验，可以与其他有类似想法的交易员合作，也可以聘请咨询顾问来协助你研究。如果策略过多，难以手工操作，应考虑提高自动化程度，除有意外情况或异常问题发生，基本就不再需要在日常交易中手动干预了。当然，你还可以选择聘请一名交易员来替你监控这些策略的执行。

与其他业务的增长不同，盈利的再投资只有数据、设备和人力上的投入，以促进交易事业的增长。当容量超过了凯利公式所建议的保守数量时，就可以开始考虑引入投资者了。这些投资者即使不愿意提供业绩费，至少也会支付设备成本。另外，你可以把你的策略（更重要的是，你的交易记录）提供给一家更大型对冲基金，并要求签署一份盈利分享合同。

最近，许多量化对冲基金遭受了巨额亏损，很多人开始质疑量化交易在长期是否可行。虽然目前讨论量化策略的死亡还为时过早，从独立交易员的角度来看，这仍是一个十分重要的问题。假如你已经实现了自动化交易并且资产呈现指数增长，是否就意味着可以放松休息、坐享财富了？回答是否定的，经验表明，随着越来越多的交易员发现这些策略并且加入竞争，这些策略就会失去盈利潜力。你必须不断研究新的策略。

每十年都会有一些突然的、重大的状态转换发生，导致某些策略的突然死亡。不管作何商业努力，经历一段高速增长期后不可避免地会进入稳定状态，行业成熟的标志便是只能获得不起眼的收益。但是，只要金融市场需要瞬时流动性，量化交易永远会有盈利基础。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

附录：MATLAB 快速回顾

MATLAB 是 Mathworks 公司开发的一款通用软件包，是许多机构量化研究员和交易员（尤其是那些从事统计套利的人员）常用的回测平台。第 3 章对 MATLAB 有过介绍，指出了这一平台相比于其他平台的优点和不足。本书大部分例子都是使用 MATLAB 编程的。许多策略都是包含上百种股票的组合交易策略，若要使用 Excel 进行回测就非常困难。这里，对 MATLAB 进行快速回顾，以帮助那些尚不熟悉这种语言的交易员，可以评估自己是否值得购买并且学习使用这一平台进行回测。

MATLAB 不仅仅是一门编程语言，它还是一个包含程序编辑器和调试器的集成开发平台，二者的操作十分简便。与 C 语言这样的传统编程语言不同，它是一种类似于 Visual Basic 的解释性语言，不需要在运行前先进行编译。由于 MATLAB 集合了大量有用的数学函数以及专门为数组（如向量或矩阵）计算而设计的数组处理语言，相比于 Excel 或 Visual Basic 在回测上更加富有弹性，功能也更强大。此外，许多在 C 或 Visual Basic 上必须使用循环的地方，在 MATLAB 中可以仅用一行代码就能实现。MATLAB 还拥有强大的文本处理功能，从而是解析、分析文本（如 Web 网页）的重要工具。此外，MATLAB 还拥有一个全面的图表库，可

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

以很容易地使用 MATLAB 绘制各种类型的图表，甚至做出动画效果（本书中的大部分图表就是用 MATLAB 绘制的）。最后，MATLAB 代码可以编译成 C 语言或 C++语言，这样，在没有安装 MATLAB 平台的电脑上这些程序也同样可以运行。事实上，有些第三方软件也可以将 MATLAB 代码转换为 C 语言源代码。

MATLAB 的基本语法与 Visual Basic 或 C 语言非常相似。例如，将数组 x 中的元素初始化：

```
x(1)=0.1;
x(2)=0.3;
% 3 elements of an array initialized.
% This is by default a row-vector
x(3)=0.2;
```

注意，这里我们不需要提前“声明”这个数组，也不需要提前告诉 MATLAB 这个数组的预期长度。如果你把程序中的“;”去掉的话，MATLAB 还会立即显示程序运行的结果，这里是对变量赋予一个值。在符号“%”之后可以写入任何注释。只要你愿意，就可以对大量元素集中进行初始化，并赋予一个相同的值：

```
% assigning the value 0.8 to all elements of a 3-vector y.
This is a row-vector. y=0.8 * ones(1,3)
```

如果你想要加总两个向量的话，可以选择老办法（C 语言中使用的方法），即使用循环：

```
for i=1:3
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
    z(i)=x(i)+y(i)% z is [0.9 1.1 1]
end
```

而 MATLAB 的强大就在于它可以十分方便地同时处理许多数组的运算，而不需要使用循环(这就是它又被称为向量处理语言的原因)。所以在 MATLAB 中仅需要输入：

```
z=x+y % z is the same [0.9 1.1 1]
```

更强大的是，MATLAB 还可以选择不同数组的一部分进行运算。你认为下面这个程序输出的结果会是什么？

```
w=x([1 3])+z([2 1])
```

x([1 3])代表了 x 向量的第一个和第三个元素，即[0.1 0.2]。z([2 1])代表了 y 向量的第二个和第一个元素，即[1.1 0.9]。所以 w 的计算结果是[1.2 1.1]。

同样可以很容易地删除部分数组元素：

```
x([1 3])=[] % this leaves x as [0.3]
```

想连接两个数组也非常容易。要在原数组之后加入行的话需要使用“;”将数组分开：

```
u=[z([1 1]);w]
% u is now
% [0.9000 0.9000;
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

```
% 1.2000 1.1000]
```

要在原数组之后加入列的话，省略“;”就可以了：

```
v=[z([1 1])w]
% v is now
% [0.9000 0.9000 1.2000 1.1000]
```

我们不仅可以通过数组的索引来选取子数组，还可以通过包含逻辑值的数组来选取子数组。比如通过逻辑数组：

```
vlogical=v<1.1
% vlogical is [1 1 0 0],where the 0s and 1's
% indicate whether that element is less than 1.1 or
% not.

vlt=v(vlogical)% vlt is [0.9 0.9]
```

事实上，我们可以用下面的方法快速获得相同的子数组：

```
vlt=v(v<1.1)% vlt is the same [0.9 0.9]
```

如果想知道数组 v 中小于 1.1 的元素的索引的话，可以使用 find 函数来实现：

```
idx=find(v<1.1);% idx is [1 2]
```

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

当然，可以反过来用这一指数数组来获取相应的子数组：

```
vlt=v(idx);% vlt is the again same [0.9 0.9]
```

到目前为止，我给出的数组例子都是一维的。当然，MATLAB同样可以处理多维数组。这里有一个二维数组的例子：

```
x=[1 2 3;4 5 6;7 8 9];
% x is
% 1 2 3
% 4 5 6
% 7 8 9
```

可以使用符号“:”来选择多维数组的所有行或列，比如：

```
xr1=x(1,:)% xr1 is the first row of x,i.e. xr1 is [1 2 3]

xc2=x(:,2)% xc2 is the second column of x,i.e. xc2 is
% 2
% 5
% 8
```

也可以使用这一方法来删除一个数组的的某一行或列：

```
x(1,:)=[] % x is now just [4 5 6;7 8 9]
```

矩阵转置的符号是“'”，MATLAB中的符号也是一样的，x的转置就是

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

x′,结果为：

```
4    7
5    8
6    9
```

数组中的元素并不一定必须是数字,也可以是字符串,甚至可以是数组本身。在 MATLAB 中这样的数组被称为单元数组。下面这个例子中,C 就是一个单元数组：

```
C={[1 2 3];['a' 'b' 'c' 'd']}
% C is
% [1 2 3]
    % 'abcd'
```

MATLAB 的一个突出优点就在于它所集成的函数几乎都能够同时应用于数组元素计算,比如：

```
log(x)% this gives
    % 1.3863 1.6094 1.7918
    % 1.9459 2.0794 2.1972
```

这种函数还有很多,比如：

sum, cumsum, diag, max, min, mean, std, corrcoef, repmat, reshape, squeeze, sort, sortrow, rand, size, length, eigs, fix, round, floor, ceil, mod, factorial, setdiff, union, intersect, ismember, unique, any, all, eval, eye,

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

ones, strmatch, regexp, regexprep, plot, hist, bar, scatter, try, catch, circshift, datestr, datenum, isempty, isfinite, isnan, islogical, randperm

如果 MATLAB 基础平台上集成的函数还无法满足需求的话，可以花钱购买 MATLAB 附加的工具箱。有些附加的工具箱对量化交易员而言是很有用的，比如最优化、偏微分方程（对衍生品交易员而言）、遗传算法、统计学、神经网络、信号处理、波、金融、金融衍生品、GARCH、金融时间序列、财务资料来源以及固定收益工具箱。要是你觉得这些工具箱过于昂贵，或还是无法满足需求的话，还可以寻找许多免费的由用户自己提供的工具，这些大多可从网上直接下载。我本书介绍过一个：由 James LeSage 开发的计量经济学工具箱（www. spatial-econometrics. com）。我之前还使用过许多其他类似的工具箱，如 Kevin Murphy 的 Bayes Net 工具箱（www. cs. ubc. cn/～murphyk/Software/BNT/bnt. html）、Kevin Sheppard 开发的 GARCH 工具箱（ww. kevinsheppard. org/research/ucsd_garch/ucsd_garch. aspx）。这些由用户自己开发的工具箱获取方便，并且有可供你咨询的数量庞大的 MATLAB 使用群体，这些都大大强化了 MATLAB 运算平台的功能。

你也可以用 MATLAB 编写自己的函数。我在本书中提供了大量的例子，这些例子都可以从我的个人网站 www. epchan. com/book 下载。事实上，将构建交易策略时常用的函数建立一个数据库是非常有必要的，随着数据库不断扩大，编写新策略的速度也会不断提升。

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

参考文献

Alexander, Carol. 2001. Market Models: A Guide to Financial Data Anal-ysis. West Sussex:John Wiley & Sons Ltd.

Chai, Soo, and Joon Lim. 2007. "Economic Turning Point Forecasting Us-ing Neural Network with Weighted Fuzzy Membership Functions." Lecture Notes in Computer Science, Springer.

Clark, Nicola. 2008. "French Bank Says Its Controls Failed for 2 Years."New York Times, Febrary 21. Available at: http://www.nytimes.com/2008/02/21/business/worldbusiness/21bank.html? ex = 1361336400&en = cf84f3776a877eac&ei=5124&partner =permalink&exprod=permalink.

Cover, Thomas. 1991. "Universal Portfolios." Mathematical Finance1 (1):1—29.

Duhigg, Charles. 2006. "Street Scene; A Smarter Computer to Pick Stock."New York Times, November 24.

Economist. 2007a. "Too Much Information." July 12. Available at: www.economist.com/finance/displaystory.cfm? story id = 9482952.

Economist. 2007b. "This Year's Model." December 13. Available at:

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

www.economist.com/finance/displaystory.cfm? story id =10286619.

Fama, Eugene, and Kenneth French. 1992. "The Cross-Section of Expected Stock Returns." Journal of Finance XLVII(2):427—465.

Fielden, Sandy. 2006. "Seasonal Surprises." Energy Risk, September. Avail-able at: www.lim.com/pdfdocs/marketing/marketfocus sept05.pdf.

Grinold, Richard, and Ronald Kahn. 1999. Active Portfolio Management. New York: McGraw-Hill.

Kaufmann, Sylvia, and Martin Scheicher. 1996. "Markov-Regime Switching in Economic Variables: Part I. Modelling, Estimating and Testing. Part II. A Selective Survey." Institute for Advanced Studies Economics Series no. 38.

Klaassen, Franc. 2002. "Improving GARCH Volatility Forecasts with Regime-Switching GARCH."Empirical Economics 27(2):363—394.

Khandani, Amir, and Lo, Andrew. 2007. "What Happened to the Quants in August 2007?" Preprint . Available at: web.mit.edu/alo/www/Papers/august07.pdf.

Lux, Hal. 2000. "The Secret World of Jim Simons." Institutional Investor Magazine, November 1.

Markoff, John. 2007. "Faster Chips Are Leaving Programmers in Their Dust." New York Times, December 17. Available at: www.nytimes.com/2007/12/17/technology/17chip.html? ex = 1355634000&en = a81769355deb7953&ei = 5124&partner = permalink&exprod=permalink.

Nielsen, Steen, and Jan Overgaard Olesen. 2000. "Regime-Switching Stock Returns and Mean Reversion." Working Paper 11 — 2000. Department of Eco-nomics, Copenhagen Business School.

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

Oldfield, Richard. 2007. Simple but Not Easy. Doddington Publishing. Poundstone, William. 2005. Fortune's Formula . New York: Hill and Wang.

Heston, Steven, and Ronnie Sadka. 2007. "Seasonality in the Cross-Section of Expected Stock Returns. "AFA 2006 Boston Meetings Paper ,July. Avail-able at: lcb1. uoregon. edu/ rcg/ seminars/ seasonal072604. pdf.

Ritter, Jay. 2003. " Behavioral Finance. " Pacific-Basin Finance Journal11(4),September:429—437.

Sharpe, William. 1994. "The Sharpe Ratio. " Journal of Portfolio Manage-ment, Fall. Available at: www. stanford. edu/ ~ wfsharpe/ art/ sr/ sr. htm.

Singal, Vijay. 2006. Beyond the Random Walk . Oxford University Press, USA.

Schiller, Robert. 2007. "Historic Turning Points in Real Estate. " Cowles Foundation Discussion Paper No. 1610. Available at: cowles. econ. yale. edu.

Schiller, Robert. 2008. "Economic View; How a Bubble Stayed under the Radar. "New York Times, March 2. Available at www. nytimes. com/ 2008/ 03/ 02/ business/ 02view. html? ex = 1362286800&en = da9e48989b6f937a&ei=5124&partner =permalink&exprod=permalink.

Taleb, Nassim. 2007. The Black Swan: The Impact of the Highly Improba-ble. Random House.

Thaler, Richard. 1994. The Winner's Curse . Princeton, NJ: Princeton Uni-versity Press. Bibliography 171

Thorp, Edward. 1997. "The Kelly Criterion in Blackjack, Sports Bet-ting, and the Stock Market. " Handbook of Asset and Liability Manage-ment, Volume I, Zenios and Ziemba (eds.). Elsevier 2006. Available at:

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

www. EdwardOThorp. com.

Toepke, Jerry. 2004. "Fill 'Er Up! Benefit from Seasonal Price Patterns in Energy Futures." Stocks, Futures and Options Magazine. March 3(3). Available at: www. sfomag. com/issuedetail. asp? MonthNameID=March&yearID =2004.

Uhlenbeck, George, and Leonard Ornstein. 1930. "On the Theory of Brow-nian Motion." Physical Review36:823—841.

Van Norden, Simon, and Huntley Schaller. 1997. "Regime Switching in Stock Market Returns." Applied Financial Economics7:177—191.

全国最低交易手续费免费办理国内正规商品股指原油期货开户 零佣金 不限资金量和交易量直接申请交易所手续费
任何地方费用比我们低 10倍赔偿差价 客服微信/QQ：958218088 期货开户和书籍下载请进:http://www.7help.net

QUANTITATIVE TRADING

ERNEST P. CHAN

How to Build Your Own Algorithmic Trading Business

“技术的进步使得发展交易策略变得更加简单。欧内斯特 · 陈使用了许多最新的量化交易技术，通过简洁描述其众多优点和少许不足，提供了一本真正服务于所有目前及即将成为交易员的专业书。”

——彼得 · 波里什（Peter Borish），计算机交易公司董事会主席兼首席执行官

“在《量化交易》一书中，欧内斯特 · 陈博士为交易策略的发展、回测检验、风险管理、编程知识、实时交易系统的研发和算法交易的运行，逐步建立起了一个最佳结构框架。”

——耶瑟 · 安瓦尔（Yaser Anwar），交易员

“量化交易是一个被神秘面纱所笼罩的极具挑战性的领域，看似艰深，只有少数精英人士才能掌握。欧内斯特 · 陈博士在其朴实无华的实践指导中，指出了进行一项成功的自动交易操作的重要基础，并分享了其在艰苦历程中获得的经验教训，为个人交易员和机构交易员指出了明确的前进方向，以避免掉进一些常见的陷阱中。”

——罗萨里奥 · M.因加尔焦拉（Rosario M. Ingargiola），Alphacet公司首席技术官

“本书为个人投资者如何建立稳定可靠的交易结构以便成功进行算法交易提供了一个有价值的视角。欧尼在构建交易系统上丰富的实践经验对那些希望交易知识更上一层楼的交易者来说是十分宝贵的。”

——雷蒙 · 卡明斯（Ramon Cummins），个人投资者

“这些年来，我所读过的关于量化交易的绝大部分书籍和文章都没有多少可取之处。在多数情况中，作者不是没有真才实学，就是因为害怕交易秘密被窃取而不愿将重要内容写在书上。欧尼有着不同的创作理念：提供了大量有价值的信息以及与量化交易各项内容有价值的广泛互动。欧尼成功地将大量琐碎艰涩的知识简化成初学者乐于接受的清晰易懂的内容。”

——史蒂夫 · 哈尔彭（Steve Halpern），HCC资本管理公司





WILEY

无防伪码者均为盗版
举报电话：(0411)84710523



ISBN 978-7-5654-1325
9 787565 413254

定价：28.00元